1-1-2018

Role Of Sirna Pathway In Epigenetic Modifications Of The Drosophila Melanogaster X Chromosome

Nikita Deshpande
Wayne State University,

ROLE OF SIRNA PATHWAY IN EPIGENETIC MODIFICATIONS OF THE DROSOPHILA MELANOGASTER X CHROMOSOME

by

NIKITA DESHPANDE

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2018

MAJOR: BIOLOGICAL SCIENCES

Approved By:

_____

Advisor Date

_____

_____

_____

DEDICATION

I would like to dedicate this dissertation to my late grandmother, Hemalata Bendre, my parents Nisha and Nitin Deshpande, my brother Nihir Deshpande and my husband, Siddharth Joshi. My grandmother has inspired me with her sheer will and open mindedness, and taught me to live my life to the fullest. My parents have inspired me to dream big and pursue my passion. My brother has always been there to hear me out and has been my greatest cheerleader. My husband has been my constant motivator and has helped me face my lowest moments. Their unconditional love and unwavering support has led to this journey's completion.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Dr. Victoria H. Meller for her exceptional mentorship. She has not only helped me become a better scientist but also taught me to be a better teacher and to be a more confident person. The effort and time she has invested in me to realize my professional goals, is truly recommendable. Without her intellectual, moral and financial support, this journey would not have been possible. I am extremely grateful for her constant guidance throughout the course of my PhD program. Her teachings and knowledge will always guide me throughout my professional career.

I thank my committee members, Dr. Lori Pile, Dr. Athar Ansari and Dr. Stephen Krawetz for their comments, advice, encouragement and criticism that has helped shaped my project. Our regular committee meetings have helped me stay focused and productive.

I have been fortunate to have the most wonderful lab mates – Dr. S. Kiran Koya, Dr. Debashish Menon, Dr. Manasi Apte, Dr. Sonal Joshi; Alissa Bults, John Butts, Reem Makki, and Maggie Sneideman. They have been a constant source of encouragement and support. I have also been fortunate to be given the opportunity to mentor undergraduates - Taania Girgla, Kassem Makki and Marah Wahbeh. I will always cherish the wonderful memories I have, learning and growing with each one of them.

I am grateful for all the valuable lessons and knowledge my teachers have bestowed upon me since my early school days. They have helped mould me into the person I am today. A special thanks to the past and present faculty and staff

at the Department of Biological Sciences. I would also like to thank the Graduate School at Wayne State University for the Summer Dissertation Fellowship, Travel Award, Graduate Enhancement Award and various career development activities.

Being 7920 miles away from family has been the biggest challenge in my PhD journey. The friends I made here, made life in Detroit fun and memorable. I have always wanted to explore the world and being in Detroit has given me a sneak peek. I made friends from all over the world, and all of them have given me a glimpse of the place they came from. I have enjoyed authentic home-cooked food, celebrated various holidays, and learned about different perspectives. This made me realize that I have taken so many things for granted, and has helped me reflect on my life and grow as an individual. I am glad to have chosen Wayne State University for my graduate studies and come to Detroit. I would like to thank my friends Dr. Shyamala Jadhav, Jasmina Kulacic, Dr. Tejeshwar Rao, Mitra Asgari, Imjoo Jang, Xiaoxiao Pu, Annisa Rochadiat and others for being there for me.

I am thankful and indebted to my parents, Nisha and Nitin Deshpande for their love, and support. Right from choosing what pre-school to enroll into, to supporting my graduate studies, they have been with me every step of the way in my professional pathway. In the process, they made innumerable sacrifices. They have supported me even through failures. My brother, Nihir Deshpande has often taken the brunt of my anxiety and kept me going through difficult times with

encouraging words and humor. I would like to acknowledge my in-laws, Bhagyashri and Chandrashekhar Joshi for their understanding, care and love.

I am left short of words to express my gratitude to my long term partner and now husband, Siddharth Joshi. He has been my constant source of inspiration, comfort, strength and encouragement. He has fought through many odds to make my life easier and has been incredibly understanding throughout my graduate life. Thank you for always being there for me!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

The DNA in the nucleus of eukaryotic cells contains genes that determine development, health and response to environmental challenges. Appropriate control of these genes is essential for each of these processes to proceed correctly. For this reason, the study of gene regulation is one of the most active and important areas of genetics. Gene regulation is usually studied in the context of local regulatory elements that control individual genes. However, eukaryotic genomes are organized into large domains of coordinated regulation. For example, the imprinted loci of mammals are clusters of genes that display regulatory patterns determined by marks placed in the parental germ line (Reik and Walter, 2001). Coordinated regulation of large domains is essential, and failure may lead to developmental abnormalities, genetic disorders, birth defects or cancer (Culbertson, 1999; Emilsson et al., 2008; Lamb et al., 2006). An extraordinary example of domain-wide regulation is modulation of sex chromosome expression, a process known as dosage compensation. In many heterogametic organisms, males have one X chromosome while females have two. In flies and mammals the X chromosome is gene-rich, while the Y chromosome has few but very important genes. Males are thus hemizygous for a large number of genes. The maintenance of a similar X to autosome expression ratio in males and females is essential for viability. Different strategies for accomplishing this have evolved independently. In mammals, females inactivate one of their X chromosomes. X inactivation is initiated and sustained by *X-inactive specific transcript* (*XIST*), a long non-coding RNA (Lee, 2009). In *C. elegans*, gene expression from each of the two hermaphrodite X chromosomes is reduced by half (Meyer, 2000). In *Drosophila*

melanogaster, males double transcribe almost all X-linked genes. Although these strategies appear dramatically different, they all are achieved by modification of chromatin on the affected chromosome (Lucchesi and Kuroda, 2015). Consequently, each system must be able to selectively identify a single chromosome. How this is achieved is poorly understood. The subject of my dissertation is a study of DNA elements that contribute to this process in flies.

Dosage compensation in *D. melanogaster* males

In *Drosophila melanogaster*, dosage compensation involves the Male Specific Lethal (MSL) complex. The MSL complex is recruited to the body of X-linked genes, where it modifies chromatin to increase transcription of the male X chromosome (Lucchesi and Kuroda, 2015). The MSL complex consists of one of two long non-coding *RNA on the X* transcripts (*roX1* or *roX2)* and five proteins, MSL1, -2, and -3, Maleless (MLE), and Males Absent on the First (MOF) (Gelbart and Kuroda, 2009; Quinn et al., 2014). Formation of the MSL complex is limited to males by the female-limited Sexlethal protein (Sxl) (Bashaw and Baker, 1997; Beckmann et al., 2005; Kelley et al., 1995). Sxl blocks MSL2 translation (Gebauer et al., 1998). MSL2 is the only male-limited protein in the MSL complex, and expression of MSL2 in the male zygote at 3 h after embryo laying (AEL) triggers formation of the intact MSL complex and X localization (Meller, 2003). MSL1 provides a scaffold for the complex through interactions with MSL2, MOF and MSL3 (Morales et al., 2004; Scott et al., 2000). MOF is a histone acetyl transferase (Hilfiker et al., 1997). The MSL complex acetylates histone 4 lysine 16, and this mark is associated with increased gene expression by enhanced transcriptional elongation (Larschan et al., 2011; Smith et al., 2000). MSL2 has been shown to have

ubiquitinating activity, but the role of this in dosage compensation remains unclear (Wu et al., 2011).

While elimination of any one of the MSL proteins is lethal to males, *roX1* and *roX2* appear fully redundant for compensation (Meller and Rattner, 2002). Loss of both *roX* transcripts results in male lethality around the time of pupation. In these males the MSL proteins are mislocalized to ectopic autosomal sites, and X-linked gene expression is reduced (Deng and Meller, 2006; Meller and Rattner, 2002). *roX1* and *roX2* are both transcribed from the X chromosome, and both have a limited ability to attract the MSL complex to active genes nearby (Kelley et al., 1999). Nevertheless, when both *roX* genes are mutated, *roX* RNA from an autosomal transgene will assemble with the MSL proteins, localize to the X chromosome and rescue males (Meller and Rattner, 2002). These observations implicate the *roX* RNAs in correct targeting of the MSL complex to the X chromosome.

How is the X selectively identified?

The MSL complex is believed to coat the X in a two-step process. Initial MSL recruitment is to Chromatin Entry Sites (CES; (Kageyama et al., 2001; Kelley et al., 1999)). These are functionally identified X-linked sites with the ability to recruit residual MSL proteins in *msl3* mutants (Fagegaltier and Baker, 2004). The MSL complex spreads from CES into nearby, active genes (Larschan et al., 2007). High resolution binding studies reveal that the MSL complex binds in the body and 3' end of actively transcribed genes (Alekseyenko et al., 2006). This pattern corresponds to the co-transcriptional H3K36me3 mark, which is bound by the MSL3 chromodomain (Larschan et al., 2011; Sural et al., 2008). Enrichment of the MSL complex and H4K16ac towards

the 3' end of genes suggests that transcriptional elongation could be facilitated, irrespective of the strength of promoter, an idea supported by gene run-on sequencing (GRO-seq) studies (Larschan et al., 2011; Smith et al., 2001). Although this model is appealing, other studies report that a modest enrichment of MSL proteins at promoters may contribute to activation of X-linked expression (Straub et al., 2013).

The CES are enriched for MSL Recognition Elements (MREs), 21 bp GA-rich motif (Alekseyenko et al., 2008; Straub et al., 2008). Chromatin-Linked Adaptor for MSL Protein (CLAMP), a zinc finger protein that binds the MRE, recruits the MSL complex by direct interaction with at least one molecule in this complex (Soruco et al., 2013). The MSL2 protein is reported to also directly interact with DNA at a subset of CES (Ramirez et al., 2015; Villa et al., 2016). Cooperation by CLAMP and MSL2 is thought to govern the properties of a subset of CES. In addition, CLAMP promotes chromatin accessibility at a distance from sites to which it is bound, and can achieve this in the absence of the MSL complex (Urban et al., 2017). Although CLAMP is a central factor in MSL complex recruitment, CLAMP binding cannot identify X chromatin. For example, CLAMP binds MREs throughout the genome, but only recruits the MSL complex to X-linked CES (Soruco et al., 2013). Indeed, MREs are only two-fold enriched on the X-chromosome (Alekseyenko et al., 2008). Additional factors must therefore contribute to X recognition.

Both *roX* genes are located on the X chromosome, and both have a limited ability to recruit the MSL complex in *cis* (reviewed in (Koya and Meller, 2011)) Additionally, both *roX* genes overlap CES (Alekseyenko et al., 2008; Kelley et al., 1999; Straub et al., 2008). However, when both *roX* genes are mutated, an autosomal *roX* transgene is able

to rescue male survival and restore dosage compensation on the X chromosome, suggesting that *roX* RNA is capable of action *in trans* to the chromatin that is modified (Meller and Rattner, 2002; Park et al., 2002). This reveals that the *roX* genes do not mark the X chromosome.

Role of siRNA pathway in X identification

The signature defect of *roX1 roX2* mutants is failure of exclusive X-chromosome recognition. A series of observations in our laboratory lead us to suspect that small RNA might cooperate with the *roX* RNAs in X recognition, and, in accord with this idea, we discovered that several genes in the siRNA pathway interact genetically with *roX1 roX2* mutants (Menon and Meller, 2009; Menon and Meller, 2012). These studies utilized the partial loss of function $roX1^{ex33}roX2\Delta$ mutant, which permits ~20% male escapers and is thus a sensitive genetic background for identification of genetic interactions. The initial study revealed that flies mutated for one copy of the endo-siRNA components *Dcr2*, *Ago2*, *D-elp1,* or *Loqs* display enhanced *roX1 roX2* lethality (Menon and Meller, 2012). Lethality was accompanied by reduced MSL localization on the X-chromosome, suggesting cooperation between siRNA and the MSL complex during identification of X chromatin. While these findings suggested that siRNA contributes to X-chromosome recognition during dosage compensation, extensive proteomic analyses of siRNA proteins and MSL complex by others have failed to find direct interactions between these pathways (Wang et al., 2013). This led us to propose that the role of siRNA in X recognition is likely to be indirect.

In *Drosophila,* siRNA processing depends on the source of RNA. Endogenous siRNAs (endo-siRNAs) are processed by Dcr-2 and R2D2, and loaded onto Ago2-

containing RNAi Induced Silencing Complex (RISC). RISC recognizes and degrades complementary mRNA. A subset of endo-siRNAs originating from structured loci are processed by Dcr-2 and the Loquacious (Loqs) isoform PD (Zhou et al., 2009). Dcr-2 associates with D-elp1, which may function in siRNA synthesis (Lipardi and Paterson, 2009). Mutations in *Dcr2, Ago2, Loqs* and *D-elp1* all enhance *roX1 roX2* male lethality, demonstrating a role for siRNA production in dosage compensation. Loss of Ago2 further reduces X-localization of the MSL proteins in a *roX1 roX2* mutant background, suggesting that siRNA might help identify X chromatin (Menon and Meller, 2012).

In considering how the siRNA pathway might promote X recognition in the fly, it may be helpful to consider how this pathway modulates chromatin in other organisms. siRNA-associated heterochromatin formation in fission yeast involves the RNAi Induced Transcriptional Silencing (RITS) complex. The RITS complex consists of Chp1 (a chromodomain protein), Ago1 (equivalent to *Drosophila* Ago2) and Tas3 (Partridge et al., 2000). siRNA bound by Ago1 recruits the RITS complex to nascent RNA, where it acts *in cis* to promote RNA interference-mediated transcriptional and post-transcriptional silencing (Sugiyama et al., 2005). Chp1 requires the methyltransferase Clr4, which deposits the H3K9 methylation mark, for localization to chromatin (Verdel et al., 2004). We hypothesize that a RITS-like complex could localize to and modify critical sequences on the fly X chromosome, and that this modification could in some way promote X recognition by the MSL complex.

Involvement of the siRNA pathway raised the question of what small RNAs were active in dosage compensation. The euchromatin of the fly X-chromosome is enriched for a clade of related 1.688^{X} repeats, also known as 1.688 g/cm^3 satellite repeats for

their density in CsCl gradients, or 359 bp repeats, the typical repeat unit length. The 1.688^{X} repeats are A-T rich and usually present in short, tandem arrays of 1 to 5 repeats. 1.688^{X} repeats at different cytological positions share an average 73% identity, but individual repeats within a cluster are near-identical. Specific clusters are denoted by a superscript denoting cytological position (Menon et al., 2014). Kuhn *et al*., (2012) noted the localization of these repeats close to or within genes, and suggested that they could play a regulatory role. The X chromosome is strikingly enriched for the 1.688^{X} repeats suggesting a potential role in dosage compensation (DiBartolomeis et al., 1992; Hsieh and Brutlag, 1979; Waring and Pollack, 1987). Interestingly, many of the 1.688^{X} repeats are transcribed, and siRNA corresponding to them has been identified in embryos (Menon et al., 2014; Usakin et al., 2007). To determine if this siRNA is active in dosage compensation, Menon *et al*. (2014) examined the effects of long single stranded RNA (ssRNA) and hairpin RNA (hpRNA) from 1.688^{X} repeats on partial loss of function *roX1 roX2* mutant males. Sense or antisense long ssRNA 1.688^{X} RNA decreased male survival by 40-70%, but hpRNA from the 1.688^{3F} repeat, which is processed into short siRNA, dramatically enhanced male survival and partially restored MSL localization on the X-chromosome (Menon et al., 2014). These findings led to the hypothesis that the siRNA pathway and the repeats on the X-chromosome are involved in X-recognition.

As the X chromosome is enriched with thousands of related 1.688^{X} repeats, as well as hundreds of CES, a level of redundancy exists that makes it impractical to study the role of an element by deletion. To determine functionality, autosomal insertions of 1.688^{X} DNA were created. These autosomal transgenes were able to recruit the MSL

complex to nearby chromatin, resulting in functional dosage compensation of nearby autosomal genes (Joshi and Meller, 2017). Compensation was enhanced by ectopic expression of cognate siRNA. This study demonstrated that the 1.688^{X} repeats are *cis*-acting regulatory sequences that help identify the X chromosome. How the 1.688^{X} repeats accomplish this remains unknown. We pursued the hypothesis that chromatin at 1.688^{X} repeats is modified by a siRNA-dependent mechanism, linking the 1.688^{X} repeats and the siRNA pathway to X-recognition.

Epigenetic modification of 1.688^{X} repeats

The objective of my dissertation was to test whether the 1.688^{X} repeats are targets of siRNA-directed chromatin modification. As no RNAi components have been found to interact directly with the MSL complex, siRNA may influence X-recognition by an indirect and novel mechanism. For example, Ago2-containing complexes could bind nascent RNAs from the X chromosome and recruit activities that alter chromatin structure or biochemistry. These modifications might, in turn, facilitate MSL recruitment and spreading into X chromatin. To explore this model, I performed a genetic screen that revealed that mutations of numerous genes encoding proteins that physically interact with Ago2 enhance the male lethality of *roX1 roX2* mutants, and thus are likely to participate in dosage compensation. This included the histone methyltransferase *Su(var)3-9*. I hypothesize that the 1.688^{X} repeats are enriched in H3K9me2 through a siRNA-dependent mechanism. I tested this by chromatin immunoprecipitation (ChIP), and found that some 1.688^{X} repeats are indeed sites of H3K9me2 enrichment, and this mark is disrupted by ectopic 1.688^{3F} siRNA production. Similar disruptions are observed in chromatin surrounding autosomal insertions of X-linked repeats. I

demonstrated that Su(var)3-9 is the enzyme that deposits the H3K9me2 mark on, and near, 1.688^{X} repeats. Finally, genes near autosomal 1.688^{X} insertions increase in expression in male larvae, and this increase is further elevated by ectopic 1.688^{3F} siRNA. These findings strongly support the hypothesis that the siRNA pathway is responsible for modifying chromatin near 1.688^{X} repeats, and that these modifications contribute to recruitment of the MSL complex. These studies are included in Chapter 3, a version of which has been submitted for publication.

Repetitive sequences have a remarkable relationship with X recognition. The MREs themselves have arisen from a mobile element that has expanded across the X chromosome (Ellison and Bachtrog, 2013). The X chromosomes of a number of closely related Drosophilids are strikingly enriched for chromosome-specific repeats, and neo-X chromosomes rapidly acquire enrichment of X-linked repeats (Gallach, 2014). In Chapter 2, I discuss the role of repeats in speciation and development of dosage compensation (Deshpande and Meller, 2014). In Chapter 4, I discuss the implications of my findings, present key questions that these studies have raised and summarize perspectives for future studies.

CHAPTER 2 SEX CHROMOSOME EVOLUTION: LIFE, DEATH AND REPETITIVE DNA

This chapter has been published as a review: Sex chromosome evolution: Life, death and repetitive DNA, Deshpande N. and Meller V.H., Fly (AUSTIN). 2014; *8*, 197-199

ABSTRACT

Dimorphic sex chromosomes create problems. Males of many species, including *Drosophila*, are heterogametic, with dissimilar X and Y chromosomes. The essential process of dosage compensation modulates the expression of X-linked genes in one sex to maintain a constant ratio of X to autosomal expression. This involves the regulation of hundreds of dissimilar genes whose only shared property is a situation close to each other on a chromosome. *Drosophila* males dosage compensate by up regulating X-linked genes two fold. This is achieved by the Male Specific Lethal (MSL) complex, which is recruited to genes on the X chromosome and modifies chromatin to increase expression. How the MSL complex is restricted to X-linked genes remains unknown. Recent studies of sex chromosome evolution have identified a central role for two types of repetitive elements in X recognition. Helitrons carrying sites that recruit the MSL complex have expanded across the X chromosome in at least one Drosophila species (Ellison and Bachtrog, 2013). Our laboratory found that siRNA from an X-linked satellite repeat promotes X recognition by a yet unknown mechanism (Menon et al., 2014). The recurring adoption of repetitive elements as X-identify elements suggests that the large and mysterious fraction of the genome called “junk” DNA is actually instrumental in the evolution of sex chromosomes.

Many eukaryotes determine sex with dimorphic sex chromosomes, such as X and Y. Y chromosomes have dramatically diminished coding potential, and this produces problems for the organisms that carry them (Charlesworth, 1996). Recombination between the X and Y produces abnormal chromosomes, and must therefore be suppressed in the male germ line. In addition, the somatic expression of X-linked genes must be adjusted so that males and females have equivalent levels of most proteins encoded on the X. Mechanisms that recognize and modulate expression from the X chromosome, termed dosage compensation, have arisen numerous times (Lucchesi and Kuroda, 2015). The diverse epigenetic machinery that has been recruited for this purpose is the subject of many excellent reviews (Lucchesi and Kuroda, 2015; Samata and Akhtar, 2018). But systems of compensation share something remarkable and less well understood: the ability to coordinate modulation of nearly all the genes on a single chromosome. We use an evolutionary perspective to argue that mobile elements and repetitive DNA are determinants of X chromosome identity in flies. New studies from our laboratory and others now implicate different types of repetitive DNA in recruitment of dosage compensation to the fly X chromosome. Interestingly, mobile elements are also a destructive force in sex chromosome evolution. The non-recombining Y chromosomes are havens for mobile DNA, leading to rapid erosion of coding potential (Rice, 1996). The duality of these roles suggests that repetitive sequences underlie the evolutionary plasticity of fly sex chromosomes.

D. melanogaster, and related species, achieves dosage compensation by increasing transcription from the male X chromosome approximately two-fold. This

occurs by selective recruitment of a ribonucleoprotein complex, the Male Specific Lethal (MSL) complex, to transcribed X-linked genes (Alekseyenko et al., 2006). The MSL complex acetylates H4 on lysine 16 (H4K16Ac), a modification that facilitates transcriptional elongation, and possibly initiation (Kind et al., 2008; Larschan et al., 2011). While the action of the MSL complex on chromatin is well studied, what limits the complex to the X chromosome remains unclear. A group of X-linked sites termed Chromatin Entry Sites (CES) recruits the MSL complex, which then moves into nearby transcribed genes (Alekseyenko et al., 2008). CES contain a 21 bp MSL Recognition Element (MRE) that binds a protein called CLAMP (Soruco et al., 2013). Knock down of CLAMP blocks X chromosome binding of MSL proteins, demonstrating its importance for X recognition. However, MREs are only modestly enriched on the X chromosome. Furthermore, CLAMP binds autosomal MREs but fails to recruit MSL proteins to autosomal sites. The question of what enables the MSL complex to selectively bind X chromatin remains open.

Comparative studies of repetitive DNA in the *Drosophila* species group reveals enrichment of different types of repetitive DNA on the X chromosome, and this occurs in parallel to the acquisition of dosage compensation. *D. miranda* provides a fascinating model as it has three X chromosomes of different ages and uses MREs to attract the MSL complex (Alekseyenko et al., 2013). The youngest X chromosomes were produced by fusions between autosomes and sex chromosomes (Steinemann et al., 1996). Orthologous to the *D. melanogaster* X is the *D. miranda* XL, over 60 million years old (Tamura et al., 2004). The *D. miranda* XR is 15 million years old, and the neo-X chromosome is 1 million years old (Bachtrog and Charlesworth, 2002). The neo-X

chromosome of *D. miranda* is in the process of acquiring MREs and enrichment for H4K16Ac in males, but this process is near-complete on the XR (Bone and Kuroda, 1996; Ellison and Bachtrog, 2013). Astonishingly, half of existing MREs on the neo-X are found in a transposable element called ISX (Ellison and Bachtrog, 2013). ISX arose by mutation of an existing helitron, and subsequent expansion of this element on the neo-X. Furthermore, some MREs on the older XR originated from a different helitron, ISXR, which also suffered a mutation that enabled MSL complex recruitment. While this is compelling, the example of *D. melanogaster* suggests that MREs are not the sole element that ensures selective recruitment of dosage compensation.

Our laboratory previously demonstrated that mutations in the siRNA (small interfering RNA) pathway are potent enhancers of mutations that impair X recognition during dosage compensation in *D. melanogaster* males (Menon and Meller, 2012). This was exciting because many organisms modify chromatin using the siRNA pathway. In brief, double stranded RNA from bidirectional transcription is processed into siRNA. siRNA associates with Argonaute proteins, which in turn guide chromatin-modifying complexes to nascent RNAs with identity to the siRNA (Ghildiyal and Zamore, 2009). However, no physical interactions between the MSL complex and components of the siRNA pathway have been discovered, suggesting an indirect mode of action. As many repetitive sequences are transcribed from both strands, these became candidates for the source of chromosome-specific siRNAs.

Our attention was attracted by a family of satellite repeats that is near-exclusive to the *D. melanogaster* X chromosome and produces siRNA. The 1.688 g/cm^3 repeats (1.688^X repeats) are dispersed throughout X euchromatin in short, tandem clusters

(Menon et al., 2014). Unusual for repetitive elements, 1.688^{X} repeats are enriched in active regions, often in introns (Kuhn et al., 2012). A role in directing dosage compensation to the X chromosome would fulfill this prediction. This inspired the suggestion that the 1.688^{X} repeats could serve to modulate expression (Kuhn et al., 2012). Examination of chromosome-specific repeats in several species revealed that X chromosome enrichment for repetitive satellites is strikingly conserved in *Drosophila* species, even when the precise sequence of these repeats is not (Gallach, 2014; Menon et al., 2014). Furthermore, the neo-X chromosome of *D. pseudoobscura* (similar to the XR chromosome of *D. miranda*) has acquired 1.688^{X} repeats, but the autosomes are devoid of them (Gallach, 2014).

Could the *D. melanogaster* 1.688^{X} repeats produce a chromosome-specific siRNA that helps identify X chromatin? To address this question, long single stranded RNA and double stranded RNA was ectopically expressed in flies with moderately reduced male survival due to impaired X recognition. Single stranded 1.688^{X} RNA further reduced male survival, but double stranded RNA from one 1.688^{X} repeat dramatically rescued males and partially restored MSL localization to the X-chromosome (Menon et al., 2014). Based on this, we put forth a model in which siRNA produced from 1.688^{X} repeats serves to recruit potential chromatin modifiers to similar X-linked regions. Rather than recruiting the MSL complex directly, we postulate that alteration of chromatin at 1.688^{X} repeats allows the X chromosome to assume a characteristic interphase conformation that facilitates recognition or distribution of the MSL complex along the chromosome. In support of this idea, the X chromosome assumes different conformations in the interphase nuclei of males and females

(Grimaud and Becker, 2009). Although our studies focused on *Drosophila*, one of the major classes of mammalian repetitive DNA has long been suspected to play a role in dosage compensation. Mammals dosage compensate by inactivating a single X chromosome in females (Disteche, 2012). Long Interspersed Nuclear Elements 1 (L1 elements) are enriched on the mammalian X and have been proposed to assist recognition of X chromatin, or spreading of silencing, during X-inactivation in mammals (Lyon, 2006). Interestingly, the formation of the inactive X territory during early differentiation is coincident with a burst of siRNA production by the L1 elements (Chow et al., 2010). We postulate that in both flies and mammals the challenge of selectively recognizing an entire chromosome is met with a combination of collaborating epigenetic pathways.

These findings raise several intriguing questions. Do X-enriched satellite repeats in other *Drosophila* species produce siRNA that promotes X recognition? If so, the rapid turnover of these repeats may be a factor in hybrid incompatibilities, which preferentially effect males, sometimes disrupting dosage compensation (Barbash, 2010). Interestingly, at least 10 Mb of pericentric X heterochromatin is composed of similar 1.688^{X} repeats in *D. melanogaster*, but absent in related species (Lohe et al., 1993). When hybrid matings introduce the *D. melanogaster* X chromosome into *D. simulans* ooplasm, the heterochromatin of the *D. melanogaster* X fails to resolve during early mitotic divisions, causing hybrid female lethality (Ferree and Barbash, 2009). One possible explanation is that *D. simulans* oocytes lack the abundant 1.688^{X} small RNAs, which may be necessary to initiate formation of pericentromeric heterochromatin at the 1.688^{X} repeats. Consistent with these ideas, removal of nearly all *D. melanogaster* X

heterochromatin by the *Zhr*^{1} translocation rescues mitosis in hybrid females (Ferree and Barbash, 2009). These studies suggest that a single, rapidly evolving class of repetitive sequences on the fly X chromosome intersects with sex chromosome biology in ways that critically influence viability and reproduction.

Eukaryotic genomes are rich with repetitive elements, often referred to as junk DNA, that have few known functions. Recent studies reveal that chromosome-specific repetitive elements and small RNA based chromatin regulation have been repeatedly adapted to guide epigenetic regulation of a chromosome. The ability to direct dosage compensation to an entire linkage group is an essential step in the evolution of dimorphic sex chromosomes. As repetitive sequences are also implicated in hybrid incompatibilities, we postulate that they confer “evolvability” not only on the predecessors of highly differentiated sex chromosomes, but also contribute to the development of species.

CHAPTER 3 CHROMATIN AT X-LINKED REPEATS THAT GUIDE DOSAGE COMPENSATION IN DROSOPHILA MELANOGASTER IS MODULATED BY THE SIRNA PATHWAY

A version of this chapter is submitted to Genetics (DESHPANDE N. and MELLER V.H., submitted)

Abstract

Many heterogametic organisms adjust sex chromosome expression to accommodate differences in gene dosage. This requires selective recruitment of regulatory factors to the modulated chromosome. How these factors are localized to a chromosome with requisite accuracy is poorly understood. *Drosophila melanogaster* males increase expression from their single X chromosome. Identification of this chromosome involves cooperation between different classes of X-identity elements. The Chromatin Entry Sites (CES) recruit a chromatin-modifying complex that spreads into nearby genes and increases expression. In addition, a family of satellite repeats that is enriched on the X chromosome, the 1.688^{X} repeats, promotes recruitment of the complex to nearby genes. The 1.688^{X} repeats and CES are dissimilar, and appear to operate through different mechanisms. Interestingly, the siRNA pathway and siRNA from a 1.688^{X} repeat also promote X recognition. We postulate that siRNA-dependent modification of 1.688^{X} chromatin contributes to recognition of nearby genes. In accord with this, we found enrichment of the siRNA effector Argonaute2 (Ago2) at some 1.688^{X} repeats. Mutations in several proteins that physically interact with Ago2, including the histone methyltransferase Su(var)3-9, enhance the lethality of males with defective X recognition. Su(var)3-9 deposits H3K9me2 on some 1.688^{X} repeats, and this mark is disrupted upon ectopic expression of 1.688^{X} siRNA. Furthermore, integration of 1.688^{X}

DNA on an autosome induces H3K9me2 deposition in nearby chromatin and enhances expression of genes on either side up to 140kb away, in a siRNA-dependent manner. Our findings are consistent with a model in which siRNA-directed modification of 1.688^{X} chromatin contributes to identification of the fly X chromosome.

Introduction

Males of many species carry one X chromosome and a gene-poor Y chromosome. Hemizygosity of the male X chromosome produces a potentially lethal imbalance in the ratio of X to autosomal gene products. This imbalance is corrected by a process known as dosage compensation, a specialized type of gene regulation that modulates expression of an entire chromosome. Different strategies to achieve dosage compensation have evolved independently. In *Drosophila melanogaster*, males increase X-linked gene expression by approximately two-fold (Lucchesi and Kuroda, 2015). This involves the activity of the Male Specific Lethal (MSL) complex. The MSL complex is recruited to active genes on the X chromosome, where it modifies chromatin to increase expression (Lucchesi and Kuroda, 2015). The MSL complex contains five proteins, Male-Specific Lethal 1, -2, and -3 (MSL1, -2, -3), Maleless (MLE), and Males Absent on the First (MOF) (reviewed in (Koya and Meller, 2011)). Enhanced transcription by the MSL complex is associated with H4K16 acetylation by MOF (Akhtar and Becker, 2000; Smith et al., 2000). H4K16 acetylation decondenses chromatin, and this may enhance transcriptional elongation of X-linked genes (Larschan et al., 2011; Shogren-Knaak et al., 2006).

The MSL complex also contains one of two non-coding *RNA on the X* (*roX1, -2*) transcripts (Quinn et al., 2014). While elimination of any one of the MSL proteins is

lethal to males, *roX1* and *roX2* are redundant for compensation. Mutation of both *roX* genes leads to mislocalization of the MSL proteins to ectopic autosomal sites in male larvae (Deng and Meller, 2006; Meller and Rattner, 2002). X-linked gene expression is reduced in these males, as is survival to adulthood. Both *roX* genes are located on the X chromosome, and both overlap Chromatin Entry Sites (CES), specialized sites with increased affinity for the MSL complex (Alekseyenko et al., 2008; Kelley et al., 1999; Straub et al., 2008).

Although much is known about the role of MSL complex in dosage compensation, how this complex selectively targets the X chromosome is poorly understood. Recognition and binding to X chromatin is believed to be a two-step process. Initial recruitment of the MSL complex to CES is followed by spreading into nearby transcribed genes (Gelbart and Kuroda, 2009). Contained within the CES are motifs called MSL Recognition Elements (MREs) (Alekseyenko et al., 2008; Straub et al., 2008). MREs are 21 bp GA-rich motifs that bind Chromatin-Linked Adaptor for MSL Protein (CLAMP), a zinc finger protein that is essential for MSL recruitment (Soruco et al., 2013). Spreading into nearby active genes is supported by interaction of MSL3 with the cotranscriptional H3K36me3 mark (Kind and Akhtar, 2007; Larschan et al., 2007; Sural et al., 2008). These mechanisms describe local recruitment of the MSL complex, but fail to explain how the MSL complex specifically targets the X-chromosome. H3K36me3 is enriched on active genes throughout the genome. MREs are only modestly enriched on the X chromosome which contains 167.7 copies of MREs per Mb compared to autosomes that contain 92.3 copies per Mb. Furthermore, CLAMP binds MREs throughout the genome, but only recruits the MSL complex to X-linked CES

(Alekseyenko et al., 2008; Soruco et al., 2013). We conclude that additional mechanisms must distinguish X and autosomal chromatin.

X-localization is disrupted in *roX1 roX2* males, making them a sensitized genetic background that can be used to identify additional factors contributing to X recognition. Using this strategy, our laboratory demonstrated a role for the siRNA pathway in recognition of the X-chromosome (Menon et al., 2014; Menon and Meller, 2012). A likely source of siRNA is a family of repeats that is near exclusive to the X chromosome. These are the AT rich, 359 bp 1.688^{X} satellite repeats, a clade of which is found in short, tandem arrays throughout X euchromatin (DiBartolomeis et al., 1992; Gallach, 2014; Hsieh and Brutlag, 1979; Waring and Pollack, 1987). Specific clusters are denoted by a superscript indicating cytological position. In support of this idea, ectopic production of siRNA from one 1.688^{X} repeat partially rescues *roX1 roX2* males (Menon et al., 2014). 1.688^{X} repeats are often close to or within genes, leading to the idea that they function as "tuning knobs" for gene regulation (Kuhn et al., 2012). In accord with these ideas, autosomal insertions of 1.688^{X} DNA enable recruitment of functional dosage compensation to nearby autosomal genes (Joshi and Meller, 2017).

The 1.688^{X} repeats share no sequence identity with the CES, and appear to act in a genetically distinct manner (Joshi and Meller, 2017). The question of how 1.688^{X} DNA promotes compensation of nearby genes is thus of great interest. We pursued the idea that siRNA-directed modifications of chromatin at 1.688^{X} repeats link the repeats and the siRNA pathway to X recognition. Reduction of the siRNA-binding effector protein Argonaute 2 (Ago2) enhances the lethality of partial loss of function *roX1 roX2* mutations, and further reduces X-localization of MSL proteins (Menon and Meller,

2012). We hypothesized that an Ago2-containing complex might localize to and modify chromatin. In accord with this idea, we find that Ago2 is enriched at 1.688^{X} repeats. Proteins interacting with Ago2 may also play a role in dosage compensation. To address this, we tested high confidence Ago2-binding proteins for genetic interactions with *roX1 roX2*, and found that mutations in several of these genes further reduced *roX1 roX2* male survival. Of particular interest is the H3K9 methyltransferase, *Su(var)3-9*, which is responsible for enrichment of H3K9me2 at a subset of 1.688^{X} repeats. H3K9me2 enrichment is disrupted upon ectopic expression of 1.688^{X} siRNA. Chromatin flanking an autosomal insertion of 1.688^{X} DNA is enriched for H3K9me2, and enrichment is enhanced by ectopic expression of 1.688^{X} siRNA. Expression of autosomal genes near the 1.688^{X} transgene is increased in male larvae, and further elevated by ectopic production of 1.688^{X} siRNA. These findings support the idea that siRNA-dependent modification of chromatin in or near 1.688^{X} repeats contributes to X recognition during dosage compensation. We propose that epigenetic modifications link the siRNA pathway, 1.688^{X} repeats on the X chromosome and X recognition.

Materials and Methods

Fly culture and Genetics

Mutations *Dcr1*^{Q1147X}, *Rm62*^{01086}, *Fmr1*$^{\Delta 113m}$, *Su(var)3-9*^{1}, *Su(var)3-9*^{2}, *smg*^{1}, *Taf11*^{1}, *Taf11*^{5}, *p53*$^{5A-1-4}$, *p53*$^{11-1B-1}$, *foxo*$^{\Delta 94}$, *PIG-S*^{e00272}, bel^{L4740}, *bel*^{6}, *barr*^{L305}, *SmD1*^{EY01516}, *vig*^{C274}, *Ago1*^{k08121}, *aub*^{QC42}, *piwi*^{06843}, *Su(var)2-10*^{2}, *egg*^{MB00702}, *G9a*^{MB11975}, P{EPgy2}^{09821}, P{EPgy2}^{15840}, *Ago2*^{414} and *FLAG.HA.Ago2* were obtained from the Bloomington Drosophila Stock Center. *Su(var)3-7*^{14} was a gift from Dr. P. Spierer (Seum et al., 2002). *ocm*^{166} was a gift from Dr. R. Kelly. *ΔDsRedΔupSET* (*upSET* in

Figure 3.4) was a gift from Dr. M. Kuroda (McElroy et al., 2017). All mutations were out crossed for five generations to minimize the effect of genetic background. Balanced stocks were constructed with outcrossed chromosomes and a laboratory reference Y-chromosome (Menon and Meller, 2009). All mutations were confirmed by phenotype or PCR. Each test scored about 1000 flies and was performed in triplicate. To express 1.688^{3F} siRNA in a *Su(var)3-9*$^{-/-}$ mutant background, we generated [*hp1.688*^{3F}] [*Sqh*-Gal4]/*In*(2LR)*Gla wg*$^{Gla-1}$*; Su(var)3-9*^{1}*/ TM3TbSb* flies and selected non-*Tb* third instar males for ChIP. The [*Sqh*-Gal4] insertion was a gift of Dr. S. Todi.

Tissue collection and chromatin preparation

Embryo collection and chromatin preparation was as previously described (Koya and Meller, 2015). Briefly, 0.5 g of 0 - 12 hr embryos were collected on molasses plates with yeast. Embryos were dechorionated for 2.5 min in bleach, crosslinked in 50 mM HEPES, 1 mM EDTA, 0.5 mM EGTA, 100 mM NaCl, 1 % formaldehyde with heptane for 20 min. Crosslinking was quenched with 125 mM glycine, 0.01 % Triton X-100, 1 X PBS for 30 min. Embryos were washed with 10 mM HEPES, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA and 0.01 % Triton X-100 and suspended in 10 mM HEPES, 1 mM EDTA, 0.5 mM EGTA, 0.1 % Na-deoxycholate and 0.02 % Na-azide for sonication in 2.5 ml buffer. Sonication was performed on ice at 35 % amplitude, 30 sec on, 59 sec off for a total time 15 min using a Fischer Scientific Model FB505 sonicator and produced 300-600 bp fragments. Chromatin was clarified by centrifuging at 13,000 rpm for 15 min, diluted 1:1 with 2 X RIPA buffer (2 % Triton X-100, 0.2 % Na-deoxycholate, 0.2 % SDS, 280 mM NaCl, 20 mM Tris-HCl pH 8.0, 2 mM EDTA, 0.02 % Na-azide, 2 mM DMSF with complete protease inhibitor (Roche)). Chromatin solution (5.5 ml) was

preabsorbed by incubation at 4˚C for 30 min with 55 µl of blocked Pierce™ Protein A agarose beads (Catalog #20333) and aliquots stored at -80˚C.

For larval chromatin, a modified protocol from (Kuzu et al., 2016) was used. 150 larvae were frozen in liquid N_2 and ground in a chilled mortar. The powder was transferred to a cooled 15 ml Dounce and homogenized with a loose pestle (10 strokes) and a tight pestle (15 strokes) in 10 ml PBS with protease inhibitor. Homogenate was made to 40 ml with PBS, crosslinked with 1 % formaldehyde for 20 min and quenched with 125 mM glycine for 30 mins. Crosslinked material was pelleted, washed once with wash buffer A (10 mM Hepes pH 7.6, 10 mM EDTA, 0.5 mM EGTA, 0.25 % Triton X-100, protease inhibitor and 0.2 mM PMSF), once with wash buffer B (10 mM Hepes pH 7.6, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1 % Triton X-100, protease inhibitor and 0.2 mM PMSF), and 3 times with TE wash buffer (10 mM Tris pH 8.0, 1 mM EDTA, 0.01 % SDS, protease inhibitor and 0.2 mM PMSF). The pellet was resuspended in 2 ml pre-RIPA buffer (0.1 % SDS, 10 mM Tris-HCl, 1 mM EDTA, protease inhibitor and 0.2 mM PMSF). Sonication was performed at settings described above for 2 min. Sonicated samples were diluted with 1 % Triton X-100, 0.1 % Na-deoxycholate, and 140 mM NaCl, centrifuged at 1500 g to clarify, aliquoted and stored at -80˚C.

Chromatin Immunoprecipitation

Seventy five micrograms of chromatin was incubated overnight with 4 µl anti-H3K9me2 (4 µg, Abcam, ab1220) or 8 µl anti-H3K9me3 (8 µg, Abcam, ab8898) at 4˚C, clarified by centrifugation at 14,000 rpm for 5 min and supernatants transferred to tubes containing 40 µl blocked Pierce™ Protein A agarose beads (Catalog #20333) and incubated 4 h at 4˚C. Washing, reverse crosslinking, DNA isolation and qPCR analysis

was as previously described (Koya and Meller, 2015). ChIP primers are presented in Appendix H.

Protein Isolation from embryos

Fifty mg of 0 - 12 hr embryos were homogenized in 250 µl RIPA buffer on ice. Homogenate was passed through a 26 gauge needle 10 - 12 times to shear DNA. Particulate matter was removed by centrifugation, and supernatant was mixed with an equal volume of 2 X SDS Sample buffer and boiled for 5 min before separation on a 15 % SDS polyacrylamide gel.

Protein blotting

Polyacrylamide gels were equilibrated in transfer buffer (48 mM Tris, 39 mM glycine, 1.3 mM SDS, 20 % methanol) for 20 min. A PVDF membrane was activated in 100 % methanol for 1 min. Filter paper and activated PVDF membranes were saturated in transfer buffer and proteins transferred using a Trans-Blot SD Semi-Dry Transfer Cell (BIO-RAD). The membrane was washed in TBST (10 mM Tris-Cl, 200 mM NaCl, 0.1 % Tween 20, pH 7.5), blocked in 5 % BSA, washed in TBST and probed overnight at 4˚ using 1:2000 mouse anti-H3K9me2 diluted in blocking solution (Abcam, ab1220) or 1:4000 goat anti-tubulin (Developmental Studies Hydrinoma Bank, E7). After washing with TBST, the membrane was incubated with alkaline phosphatase conjugated secondary antibodies (goat anti-mouse, Sigma, A3562 or rabbit anti-goat, Sigma, A4062), washed and developed in 100 mM diethanolamine, 100 mM NaCl, 5 mM $MgCl_2$, pH 9.5 containing 33 µg/ml Nitroblue Tetrazolium (NBT) and 165 µg/ml 5-Bromo-4-chloro-3-indolyl phosphate (BCIP). Signals were quantified by ImageJ.

Quantitative RT-PCR

Total RNA was isolated from 50 third instar male larvae or 100 mg dechorionated embryos using Trizol reagent (Invitrogen) as previously described (Koya and Meller, 2015). One microgram of RNA was reverse-transcribed using random hexamers and ImProm-II reverse transcriptase (Promega). Duplicate reactions were amplified using iTaq Universal SYBR Green Supermix (Bio-Rad) with an Mx3000P Real-Time PCR system (Stratagene). Primers are in Appendix H. Values were normalized to *dmn* (DCTN2-p50) and expression calculated using the efficiency corrected comparative quantification method (Pfaffl, 2001).

Results

Ago2 localizes at 1.688^{X} repeats.

Ago2 localization was determined using a FLAG-tagged *Ago2* transgene that was first tested for rescue of the dosage compensation function of Ago2. Males with the partial loss of function *roX1*^{ex40}*roX2Δ* chromosome have high survival, as do *Ago2*$^{-/-}$ flies, but synthetic lethality is observed in *roX1*^{ex40}*roX2Δ*/Y; *Ago2*$^{-/-}$ males (Menon and Meller, 2012). One copy of a FLAG-*Ago2* transgene rescues these males, demonstrating that the FLAG tag does not disrupt the dosage compensation function of Ago2 (Figure 3.1, 3.2). Chromatin from FLAG-*Ago2*; *Ago*$^{-/-}$ embryos, and from a reference strain lacking the FLAG-*Ago2* transgene, was immunoprecipitated with anti-FLAG antibodies and enrichment determined by quantitative PCR (qPCR). FLAG-Ago2 was enriched at the *Hsp70* promoter, a site known to bind Ago2 (Cernilogar et al., 2011) (Figure 3.3 A). In contrast, a control region in the *dmn* gene displayed no enrichment. We then examined FLAG-Ago2 enrichment at a panel of six representative 1.688^{X} repeats that differ in location, copy number, sequence, genetic environment and

transcription level (Table 3.1). Interestingly, five of these show enrichment of FLAG-Ago2 over the repeats, but little or no enrichment in flanking regions (Figure 3.3 B). We conclude that Ago2 localizes at many 1.688^{X} repeats, a finding that is consistent with involvement of Ago2 in siRNA-directed recruitment of chromatin modification at or around these regions.

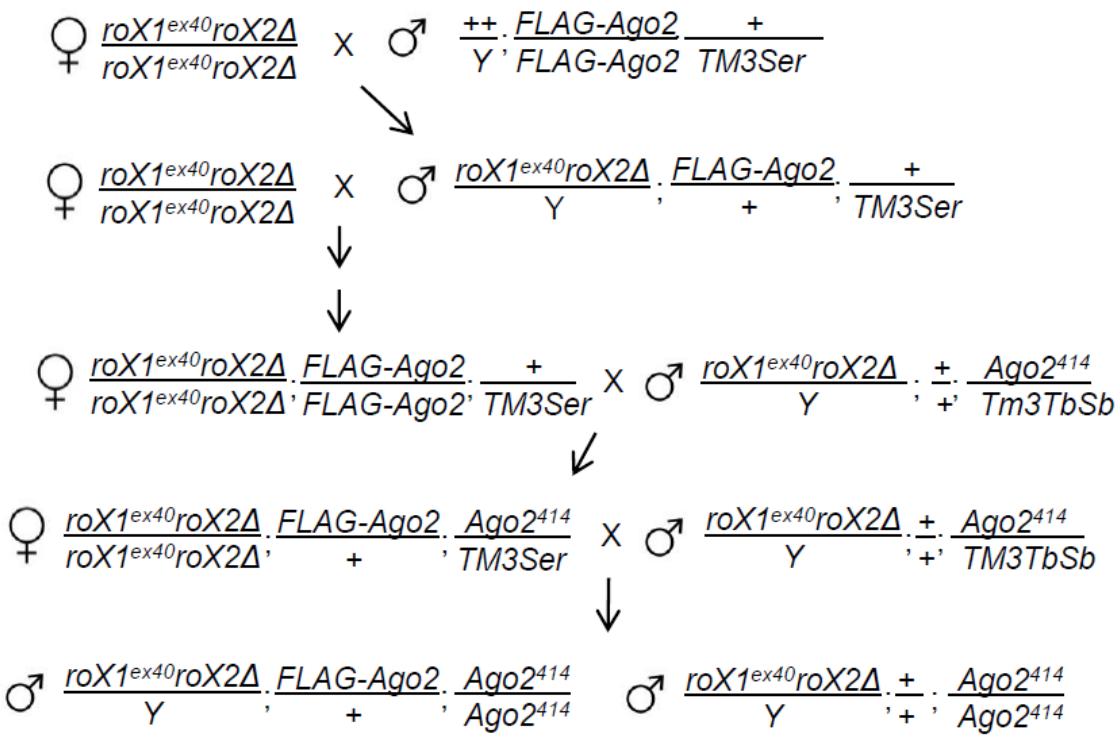


Figure 3.1. Mating scheme to express FLAG-Ago2 in Ago2 mutants. *FLAG-Ago2* is marked by w^+, enabling identification through the multiple crossing steps. *roX2Δ* is also marked by w^+. Presence of both, *roX2Δ* and *FLAG-Ago2* results in a darker red eye color.

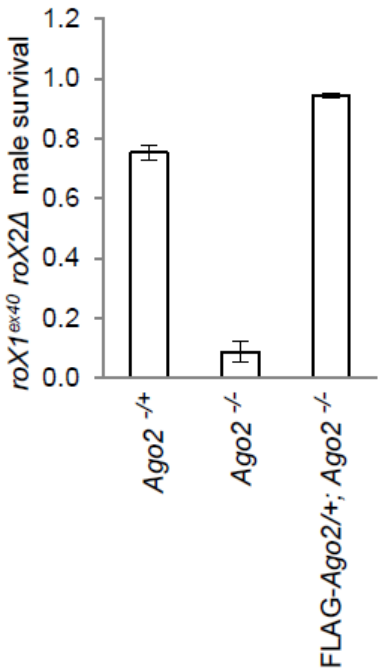


Figure.3.2. FLAG-Ago2 rescues the Ago2 dosage compensation function. Ratio of test male survival to control male survival is plotted. A FLAG-*Ago2* transgene (right) rescues the synthetic lethality of *roX1*^{ex40} *roX2Δ*/Y; *Ago2*$^{414/414}$ males (center).

Repeat		1.688[1A]	1.688[3C]	1.688[3F]	1.688[4A]	1.688[7E]	1.688[7F]
Cytological position		1A1	3B5	3F3	4A4	7E5	7F3
	Scaffold coordinates on X	204,046..241,257	2,768,669..2,770,136	3,857,647..3,858,186	4,070,631..4,071,316	8,369,270..8,369,783	8,530,749..8,531,260
Copy number	Genomic Scaffold	98	3	2.5	2	1.5	1.5
	Laboratory reference	-	3.5	3.5	2.5	2.5	1.5
Repeat properties	Percent match to 1.688[3F]	89	69	100	95	71	69
	Genomic environment	37kb of repeats between convergent genes	Intronic	Between convergent genes	Isolated	Isolated	Intronic
Transcription	EST abundance	0	2	9	2	0	26
	qRT PCR[a]	0.00026	0.01	0.3	0.1	0.04	1.5
	RNA polII enrichment	-	-	+	-	-	-

Table 3.1. Panel of 1.688^{X} repeats used in this study. Cytological positions of 1.688^{X} repeats and scaffold coordinates were determined from Flybase (Release 6). The size of some repeat clusters in our laboratory reference strain was found to differ from the genomic scaffold. See Appendix D for determination of copy number. Similarity to 1.688^{3F} was determined by BLAST. EST abundance was inferred from assigned ESTs in Flybase. RNA polII enrichment is derived from ChIP-seq of 6-8 h mesoderm (Monfort and Furlong, 2015). [a] Quantitative RT-PCR (qRT PCR) is normalized to repeat copy number (see Figure 3.11).

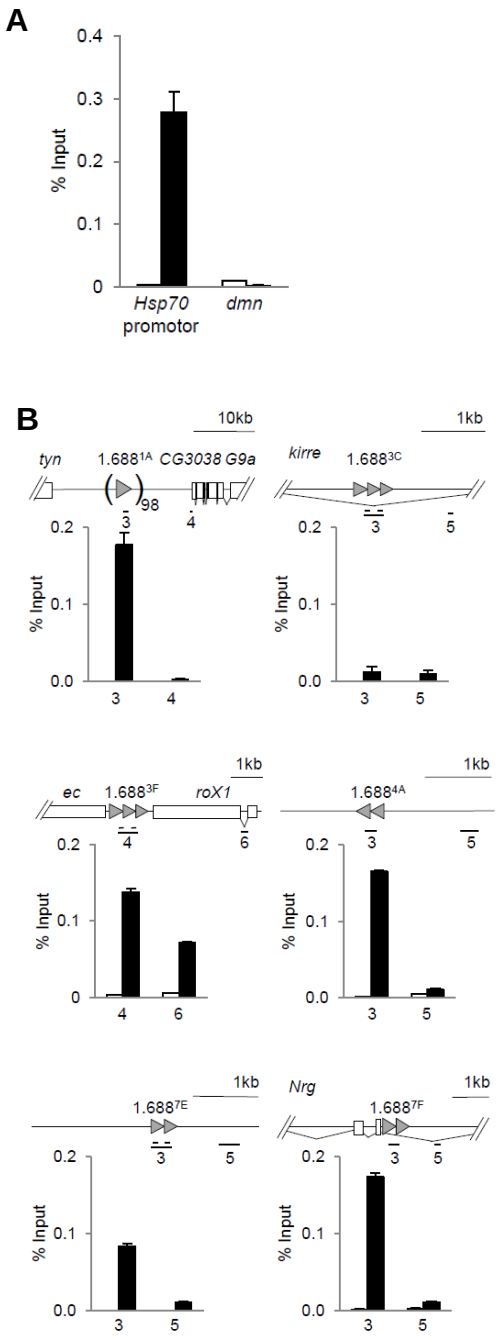


Figure 3.3. FLAG-Ago2 localizes at 1.688^{X} repeats. Chromatin from the laboratory reference strain (white) and *Ago2*$^{414/414}$; FLAG-*Ago2* (black) embryos was precipitated with anti-FLAG antibody. Enrichment normalized to input is shown. (A) The *Hsp70* promoter displays enrichment, but a control region in *dmn* does not. (B) FLAG-Ago2 enrichment is detected at several 1.688^{X} repeats (gray arrowheads). Approximately 100 copies of the 1.688^{1A} repeats are situated between *tyn* and CG3038. The 1.688^{3C} repeats are within a large *kirre* intron (splicing indicated by diagonal lines). Primers used for analysis are indexed by gene and amplicon number and presented in Appendix H. Standard error of two biological replicates is shown.

Genetic interactions identify an Ago2-interaction network that participates in dosage compensation.

Argonaute proteins in the RNA Induced Transcriptional Silencing (RITS) complexes of *S. pombe* and plants recruit chromatin modifiers to nascent transcripts (reviewed in (Meller et al., 2015)). To explore the possibility of Ago2-interacting proteins participating in X chromosome recognition, we screened genes in an Ago2-interaction network for genetic interaction with *roX1 roX2*. A map of high probability Ago2-interactors was created using BioGRID (Stark et al., 2006), and esyN (Bean et al., 2014) (Figure 3.4 A; see Appendix E for inclusion criteria). Members of this network were examined for genetic interactions with the partial loss of function *roX1*^{ex33}*roX2Δ* X chromosome. *roX1*^{ex33}*roX2Δ* males display partial mislocalization of MSL proteins and eclose at 20 % of normal levels (Deng et al., 2005b). Reduction of proteins that participate in X recognition further disrupts X localization and enhances *roX1*^{ex33}*roX2Δ* male lethality (Menon and Meller, 2012). Females are fully viable and fertile when the *roX* genes are mutated. *roX1*^{ex33}*roX2Δ* females were mated to males that were heterozygous for a mutation in the gene being tested (Figure 3.5 A). All sons are *roX1*^{ex33}*roX2Δ*/Y, and heterozygous (experimental) or wild type (control) for the gene of interest. Normalized survival (experimental /control) reveals enhancement of *roX1 roX2* male lethality (Figure 3.4 B, C). Daughters, which do not dosage compensate and are heterozygous for *roX1*^{ex33}*roX2Δ,* do not display altered survival upon mutation of Ago2-interacting genes. As *G9a* is located on the X chromosome, a modified strategy to test this gene is presented in Figure 3.5 B, C.

Normalized survival of *roX1*^{ex33}*roX2Δ* males with mutations in the Ago2-interaction network is presented in Figure 3.4 B. Genes displaying significant interactions are noted by pink symbols, and those showing no interaction are blue in Figure 3.4 A. We confirmed a previously identified siRNA-processing sub-network containing *Dcr2*, *Elp1*, and *loqs* (Figure 3.4 A, dotted line; (Menon and Meller, 2012)). The present study identified several additional Ago2-interactors, including a potential chromatin-modifying sub-network containing *Dcr1, Fmr1, Rm62,* and the histone methyltransferase *Su(var)3-9* (green, Figure 3.4 A). Su(var)3-9 deposits H3K9me2 and acts with Rm62 to re-silence active chromatin (Boeke et al., 2011).

Additional chromatin modifiers and genes in other small RNA pathways were also tested (Figure 3.4 C). A previous study found no interaction between *roX1*^{ex33}*roX2Δ* and the piRNA pathway genes *aub* and *piwi,* or the miRNA pathway gene *Ago1*, a finding replicated here. Since our findings point towards involvement of chromatin modifiers, we tested the chromatin regulatory factor *Su(var)2-10* and two additional H3K9 methyltransferases, *eggless* (*egg*) and *G9a* (Figure 3.4 C). None of these modified *roX1*^{ex33}*roX2Δ* survival. Mutations in *Su(var)3-7*, important for heterochromatin formation, and *upSET*, which maintains heterochromatin and H3K9me2 levels, enhance *roX1*^{ex33}*roX2Δ* male lethality (McElroy et al., 2017; Spierer et al., 2008). *Over compensating males* (*ocm)* has an unusual dosage compensation phenotype as mutations in *ocm* rescue males with insufficient MSL activity, suggesting that it might act as a governor of activation (Lim and Kelley, 2013). Interestingly, mutation of *ocm* significantly increased the survival of *roX1*^{ex33}*roX2Δ* males, supporting the idea that *ocm* normally restrains activation. The P{EPgy2}^{09821} and P{EPgy2}^{15840} strains, used to

outcross *Su(var)3-9* and *barr* mutants, display no interaction and serve as controls for genetic background. Taken together, these findings suggest that several genes that deposit H3K9me2, maintain this mark or participate in heterochromatin formation also contribute to X chromosome dosage compensation. At first glance these observations appear to be at odds with X chromosome hypertranscription, the ultimate consequence of X chromosome recognition.

Ectopically expressed 1.688^{3F} siRNA disrupts H3K9me2 patterns

Previous studies found that ectopically produced 1.688^{3F} siRNA partially rescues *roX1 roX2* males and increases X localization of the MSL complex (Menon et al., 2014). The mechanism by which siRNA promotes X recognition is unknown. The discovery that insertion of 1.688^{X} DNA on an autosome enables functional compensation of nearby genes, and the enhancement of this effect by ectopic 1.688^{3F} siRNA, suggests siRNA action through cognate genomic regions (Joshi and Meller, 2017). In accord with this idea, an autosomal *roX1* transgene also enables compensation of nearby genes, but is unaffected by 1.688^{3F} siRNA. To test the idea that 1.688^{3F} siRNA directs epigenetic modification of 1.688^{X} chromatin, we used ChIP to analyze chromatin around 1.688^{X} repeats on the X chromosome. ChIP-qPCR detected H3K9me2 enrichment in 4 out of 6 repeats (white bars, Figure 3.6). As H3K9me2 enrichment was not uniform, we considered additional factors that might determine this mark, and noted that only repeats showing evidence of transcription were enriched for H3K9me2, consistent with the idea of Ago2-dependent recruitment to nascent transcripts (Verdel et al., 2004). Upon ectopic expression of 1.688^{3F} siRNA a dramatic disruption of H3K9me2 was observed in and around 1.688^{X} repeats (black bars, Figure 3.6). For example, 1.688^{3F}

and 1.688^{4A} display peaks of H3K9me2 in wild type flies, but this mark was reduced over the repeats and increased in surrounding regions by elevated 1.688^{3F} siRNA. Untranscribed repeat clusters at 1.688^{1A} and 1.688^{7E} show no H3K9me2 enrichment in wild type flies, but gained H3K9me2 upon expression of 1.688^{3F} siRNA. In contrast, no enrichment of H3K9me3 in or near 1.688^{X} repeats was detected in wild type or 1.688^{3F} siRNA-expressing embryos (Figure 3.7). This is as expected as H3K9me2 is found in facultative heterochromatin by contrast H3K9me3 is found in constitutive heterochromatin such as at telomeres and centromeres (Becker et al., 2016). We conclude that some 1.688^{X} repeats are enriched for H3K9me2, and that ectopic production of cognate siRNA broadly disrupts this mark.

Su(var)3-9 deposits H3K9me2 at 1.688^{X} repeats

The identification of Su(var)3-9 as an indirect binding partner of Ago2, observation of a genetic interaction between *Su(var)3-9* and *roX1 roX2* and enrichment of H3K9me2 on some 1.688^{X} repeats suggests that Su(var)3-9 could be modifying 1.688^{X} repeats. *D. melanogaster* has three histone H3K9 methyltransferase, *Su(var)3-9*, *eggless*, and *G9a*, but only *Su(var)3-9* mutations enhance the male lethality of *roX1 roX2* ((Swaminathan et al., 2012); Figure 3.4). To determine if Su(var)3-9 is responsible for H3K9me2 at 1.688^{X} chromatin, 3rd instar larvae mutated for *Su(var)3-9,* or mutated for *Su(var)3-9* and expressing 1.688^{3F} siRNA, were generated (Figure 3.8). H3K9me2 enrichment is virtually eliminated over 1.688^{X} repeats in *Su(var)3-9*$^{-/-}$ mutants (gray bars, Figure 3.9) and remains low in Su(var)3-9$^{-/-}$ larvae that express 1.688^{3F} siRNA (black bars, Figure 3.9). This reveals that Su(var)3-9 deposits H3K9me2 at 1.688^{X} chromatin in wild type flies, and eliminates the possibility that a different methyltransferase is recruited to these

regions following ectopic expression of 1.688^{3F} siRNA. Disruption of H3K9me2 upon expression of 1.688^{3F} siRNA thus reflects changes in the localization or activity of Su(var)3-9.

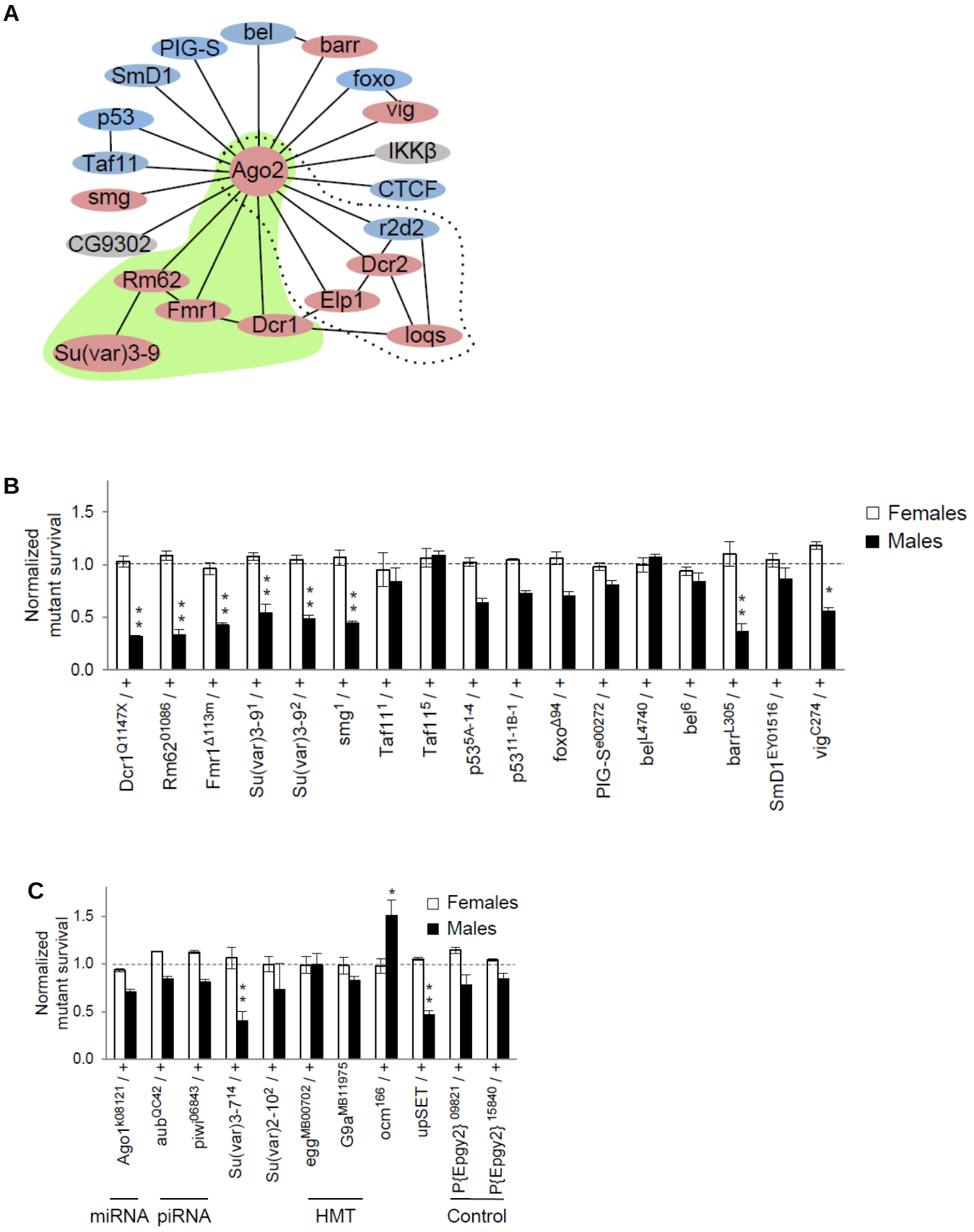
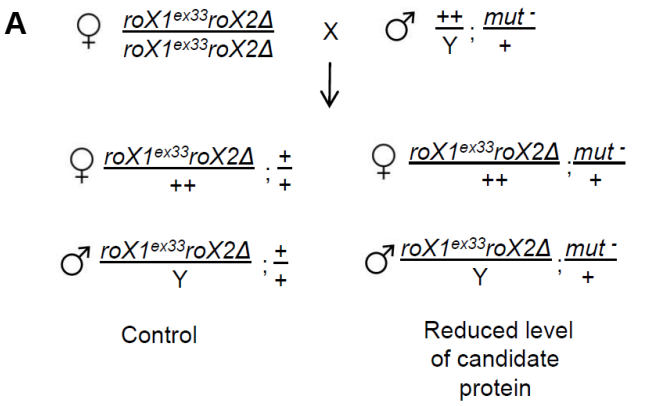
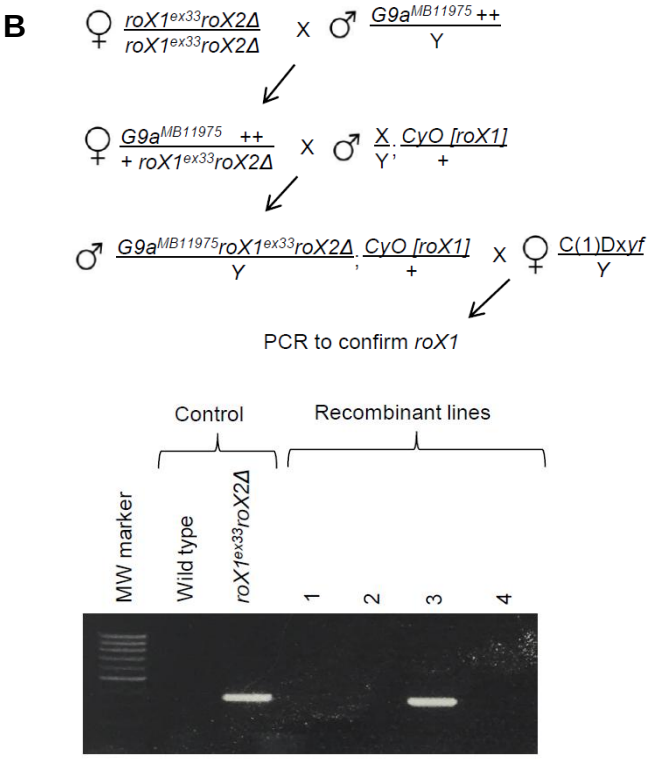


Figure 3.4. Ago2-interactors participate in dosage compensation. (A) Map of Ago2-interacting proteins. Genes displaying a genetic interaction with *roX1*^{ex33}*roX2Δ* are pink, and those for which a significant interaction has not been detected are blue. Genes in gray are untested. A previously reported siRNA-production sub-network is highlighted by the dotted line. A putative chromatin-modifying sub-network identified in the present study is highlighted in green. Well-curated, high probability interactions from BioGRID and esyN are depicted by solid lines. See Appendix E for inclusion criteria. (B) Mutations in many Ago2-interacting proteins reduce the recovery of *roX1*^{ex33}*roX2Δ* males (black; *roX1*^{ex33}*roX2Δ/Y; mut/+* normalized to *roX1*^{ex33}*roX2Δ/Y; +/+*). Females are unaffected (white; *roX1*^{ex33}*roX2Δ/++; mut/+* normalized to *roX1*^{ex33}*roX2Δ/++; +/+*). (C) Additional controls and genes of interest. The mating strategy to test X-linked G9a is presented in Figure 3.5 C. See Materials and Methods for *upSET* description. SEM is represented by error bars. Significance of ≤0.05 (*) and ≤0.001 (**) was determined using the Student's two sample *t*-test.





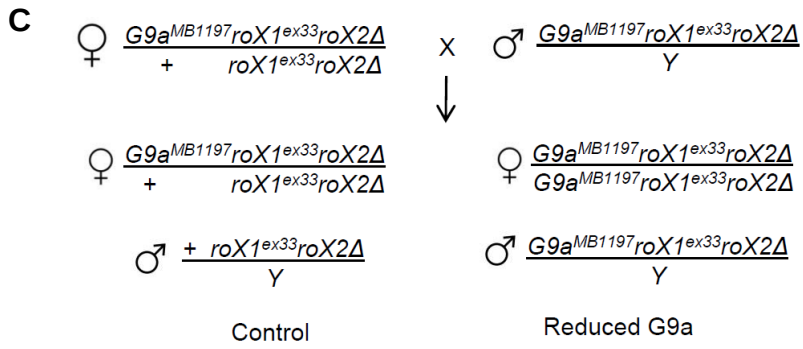


Figure 3.5. Detection of genetic interactions between *roX1 roX2* and candidate genes. (A) *roX1*^{ex33}*roX2Δ* females were mated to males heterozygous for a mutation in the gene of interest. The survival of sons mutated for the gene of interest (bottom right) is divided by that of control brothers (bottom left) and presented in Figure 3.4 B and C. In an otherwise wild type background, *roX1*^{ex33}*roX2Δ* allows 20 % adult male escapers. Females do not dosage compensate and serve as an internal control. (B) Mating scheme to generate *G9a roX1*^{ex33}*roX2Δ* mutants. *G9a* is marked by GFP and *roX2Δ* is marked by w^+, enabling identification of recombinants carrying both mutations. Recombinants that were also *roX1*^{ex33}, predicted to be 33.5% of the lines screened, were identified by PCR. *G9a roX1*^{ex33}*roX2Δ* recombinant 3 was used in subsequent studies. (C) Testing for genetic interaction between G9a and *roX1*^{ex33}*roX2Δ*. Heterozygous *G9a*^{MB1197} *roX1*^{ex33}*roX2Δ*/+ *roX1*^{ex33}*roX2Δ* females were mated to *G9a* ^{MB1197} *roX1 roX2* males. *G9a*^{MB1197} is marked with EGFP. The survival of *G9a* ^{MB1197} *roX1*^{ex33}*roX2Δ* sons (EGFP-positive, right) was divided by that of EGFP-negative + *roX1*^{ex33}*roX2Δ* sons (left). EGFP intensity differentiates F1 females that are homozygous or heterozygous for *G9a* ^{MB1197}.

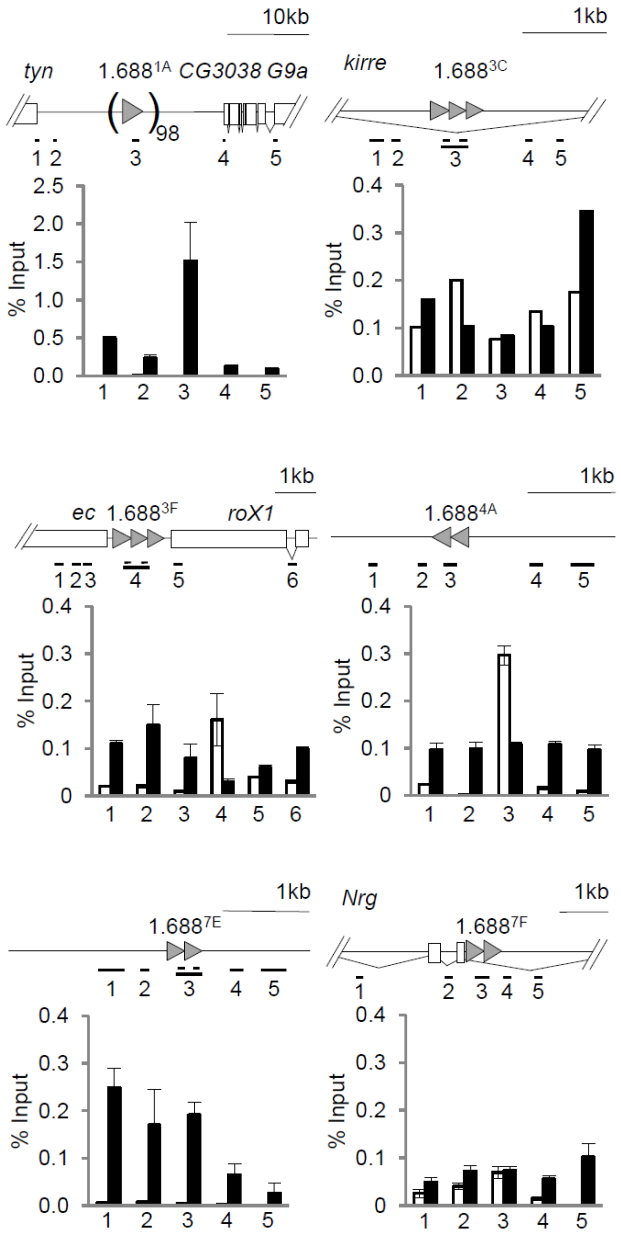


Figure 3.6. Elevated 1.688^{3F} siRNA disrupts H3K9me2 enrichment around 1.688^{X} repeats. Chromatin from wild type embryos (white) and embryos ectopically producing 1.688^{3F} siRNA (black) was immunoprecipitated with antibody to H3K9me2. Enrichment over input was determined by quantitative PCR (qPCR). The standard error of two biological replicates is shown. Primers used for analysis are presented in Appendix H.

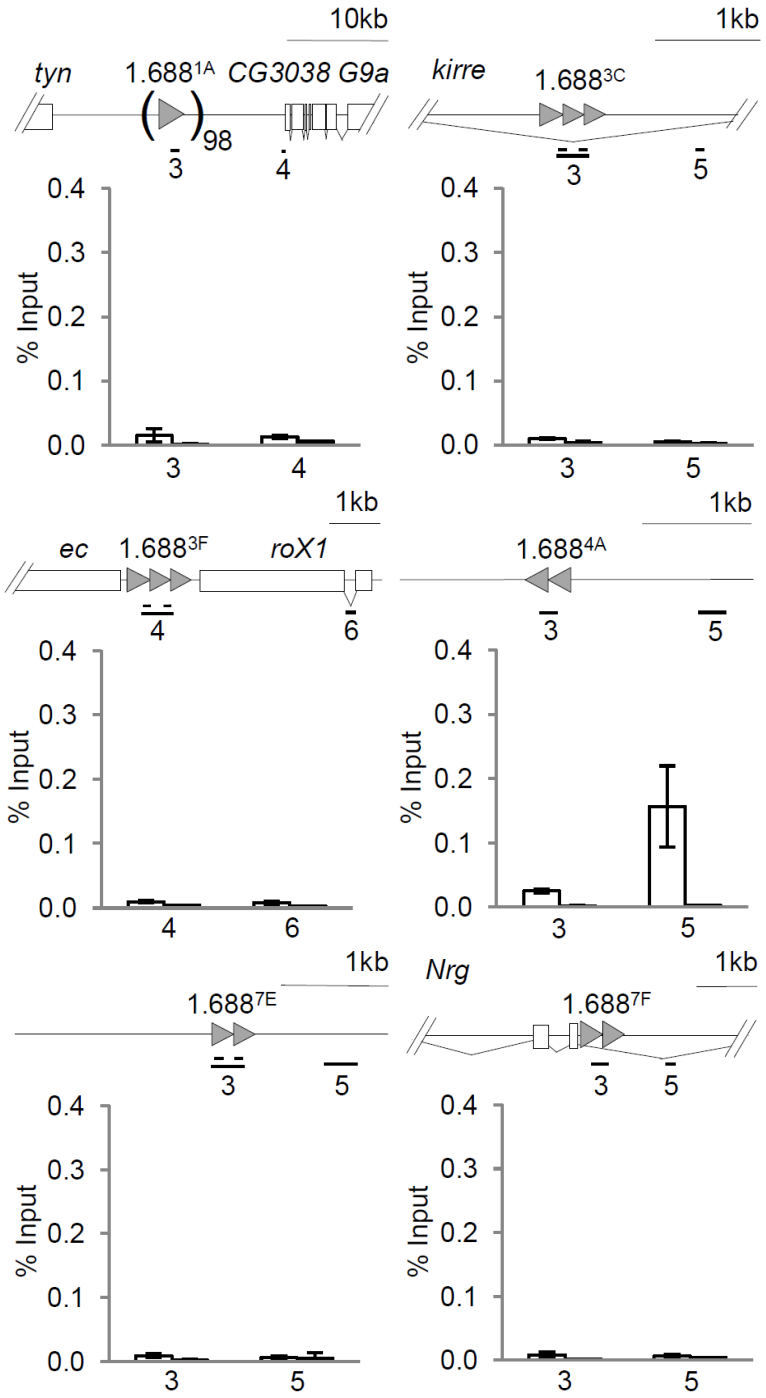


Figure 3.7. H3K9me3 is not enriched over 1.688^{X} repeats or altered by ectopic expression of 1.688^{3F} siRNA. Chromatin from 6-12 hr embryos was immunoprecipitated with antibody to H3K9me3. DNA was analyzed by qPCR using primers within 1.688^{X} repeats (gray triangles) or in flanking regions. Approximately 100 copies of 1.688^{1A} are present between *tyn* and *CG3038*. Primers indexed by gene and amplicon number are presented in Appendix H. No significant enrichment within repeats, or change in H3K9me3 within repeats, is observed following siRNA expression. Standard error is derived from two biological replicates.

To determine how far from 1.688^{X} repeats the H3K9me2 disruption extends, regions 10-26 kb from repeats were examined. In each case, these regions displayed increased H3K9me2 in embryos with ectopic 1.688^{3F} siRNA expression (Figure 3.10 A). This suggested the possibility of a global change in H3K9me2. To address this possibility we probed protein blots from wild type and 1.688^{3F} siRNA-expressing embryos to determine the levels of this modification. In spite of apparently wide-spread elevation of H3K9me2, no evidence for a global change in H3K9me2 level is detected (Figure 3.10 B). As most H3K9me2 is found in heterochromatic regions that comprise ~30% of the fly genome, changes in euchromatic regions may represent a negligible portion of the nuclear pool.

H3K9me2 is generally thought to be repressive, but compensation in flies occurs by increased expression of X-linked genes. To determine if changes in H3K9me2 enrichment correlate with changes in transcription, expression of genes near 1.688^{X} repeats was examined in wild type and 1.688^{3F} siRNA-expressing embryos. Consistent with H3K9me2 having a repressive effect, 1.688^{3F} siRNA decreases accumulation of RNA from non-coding intragenic or intronic regions with elevated H3K9me2 (Figure 3.11). The apparent increase in 1.688^{3F} expression (Figure 3.11) is from the transgene used to produce ectopic 1.688^{3F} siRNA. We detected dramatic reductions in messages adjacent to 1.688^{1A} (*tyn*, *G9a*) and 1.688^{3F} (*ec*, *roX1*). In spite of a 90% reduction in *ec* transcript in embryos expressing 1.688^{3F} siRNA, adults of this genotype do not display the rough eye *ec* phenotype. It is possible that ectopic 1.688^{3F} siRNA has a more pronounced effect in embryos, whose undifferentiated cells may be particularly susceptible to chromatin-based disruption. Mature patterns of chromatin organization

are established by late larval life, and these may be more resistant. To test this, we examined expression in wild type and 1.688^{3F} siRNA-expressing 3rd instar male larvae, and found that *tyn*, *G9a* and *ec* regained wild type levels of transcript, and *roX1* was also largely restored (Figure 3.11). This might be due to recovery upon establishment of mature chromatin. The precise reason for the differences between embryos and larvae are uncertain, but restoration of normal gene expression by the 3rd instar larvae is consistent with the lack of phenotype in otherwise wild type flies that ectopically express 1.688^{3F} siRNA (Menon et al., 2014).

The discovery that animal age influenced the response to ectopic siRNA prompted us to determine the time point at which H3K9me2 is established at 1.688^{X} repeats. A possible scenario is that this mark is placed before MSL localization, and acts in some way to guide X recognition. X-localization of the MSL complex occurs at 3 hr after egg laying (AEL) (Meller, 2003; Rastelli et al., 1995). We determined H3K9me2 enrichment at 1.688^{3F} in embryos before the MSL complex binds to the X (1.5-3 hr), during initial MSL recruitment (3-4 hr), and at 4-6 hr and 6-12 hr. In contrast to our prediction, H3K9me2 is first detected on 1.688^{3F} between 6 and 12 h AEL, after X localization of the MSL complex has occurred (Figure 3.12). We conclude that H3K9me2 at 1.688^{X} repeats is unlikely to guide initial X recognition, but may serve at a later time point to facilitate spreading of this mark or enforce the stability of X recognition.

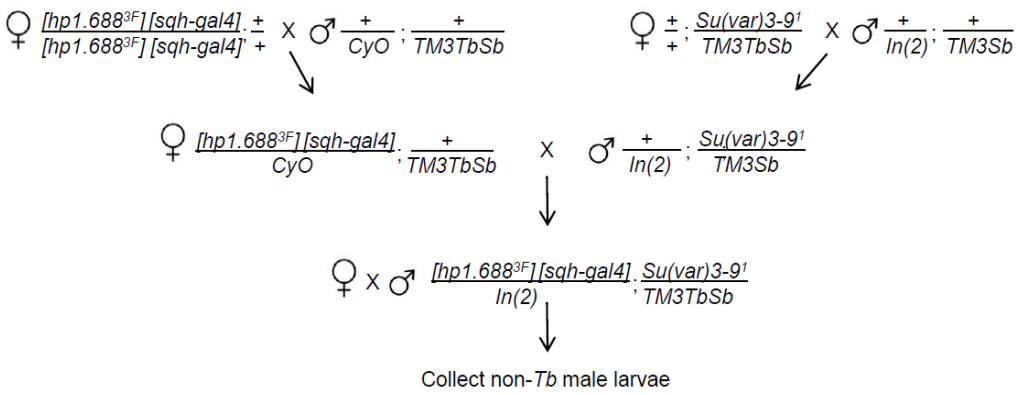


Figure 3.8. Mating scheme to generate *Su(var)3-9* mutants expressing 1.688^{3F} siRNA. [*hp1.688^{3F}*] [*Sqh*-Gal4] is marked by w^+, enabling identification through the multiple crossing steps. *Su(var)3-9^1* was followed by a 3rd chromosome balancer and at the final step non-*Tb* 3rd instar males were collected for ChIP.

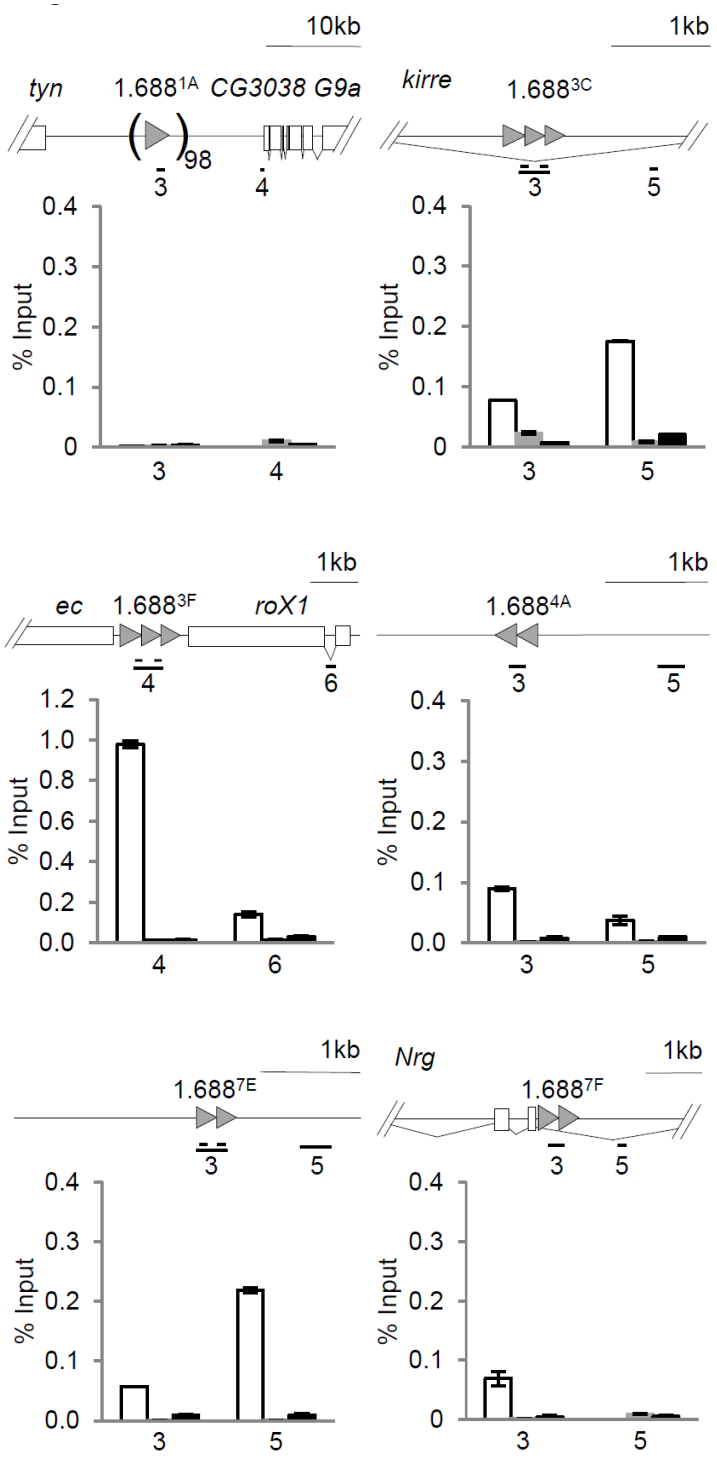
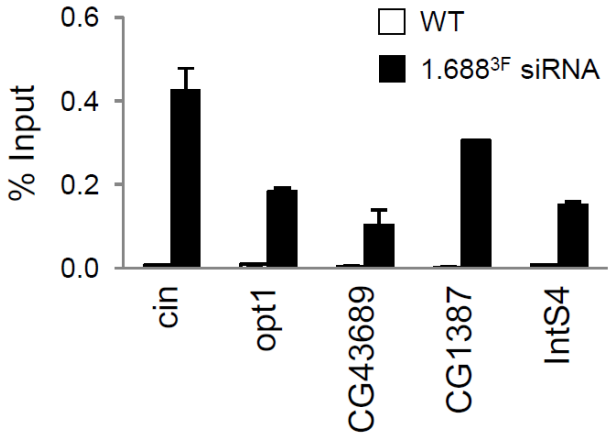


Figure 3.9. Su(var)3-9 deposits H3K9me2 at some 1.688^{X} repeats. Chromatin from wild type male larvae (white), *Su(var)3-9^{1}/Su(var)3-9^{1}* male larvae (gray), and *Su(var)3-9^{1}/Su(var)3-9^{1}* males ectopically expressing 1.688^{3F} siRNA (black) was immunoprecipitated with antibody to H3K9me2. Enrichment normalized to input is shown. Standard error is derived from two biological replicates. See materials and methods and Figure 3.8 for full genotypes and larval selection strategy.

A

Gene	Distance from repeat
cin	15 kb from 1.688^{1A}
opt1	10 kb from 1.688^{3F}
CG43689	10 kb from 1.688^{4A}
CG1387	14 kb from 1.688^{7E}
IntS4	26 kb from 1.688^{7F}



B

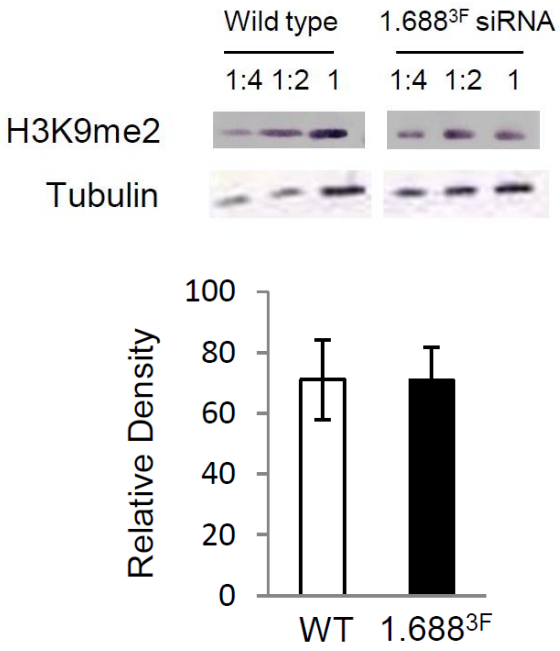


Figure 3.10. Widespread alteration in H3K9me2 around 1.688^{X} repeats is not reflected in global H3K9me2 level. (A) Genes over 20 kb from 1.688^{X} repeats display increased H3K9me2 following ectopic 1.688^{3F} siRNA production. (B) Western blot of histones from control (wild type) and 1.688^{3F} siRNA-expressing embryos does not detect a change in H3K9me2 level. H3K9me2 levels in 6-12 h embryos were compared to a tubulin loading control. Sample dilution was used to confirm signal linearity. Signal intensity was determined by ImageJ software. Standard error is derived from three biological replicates.

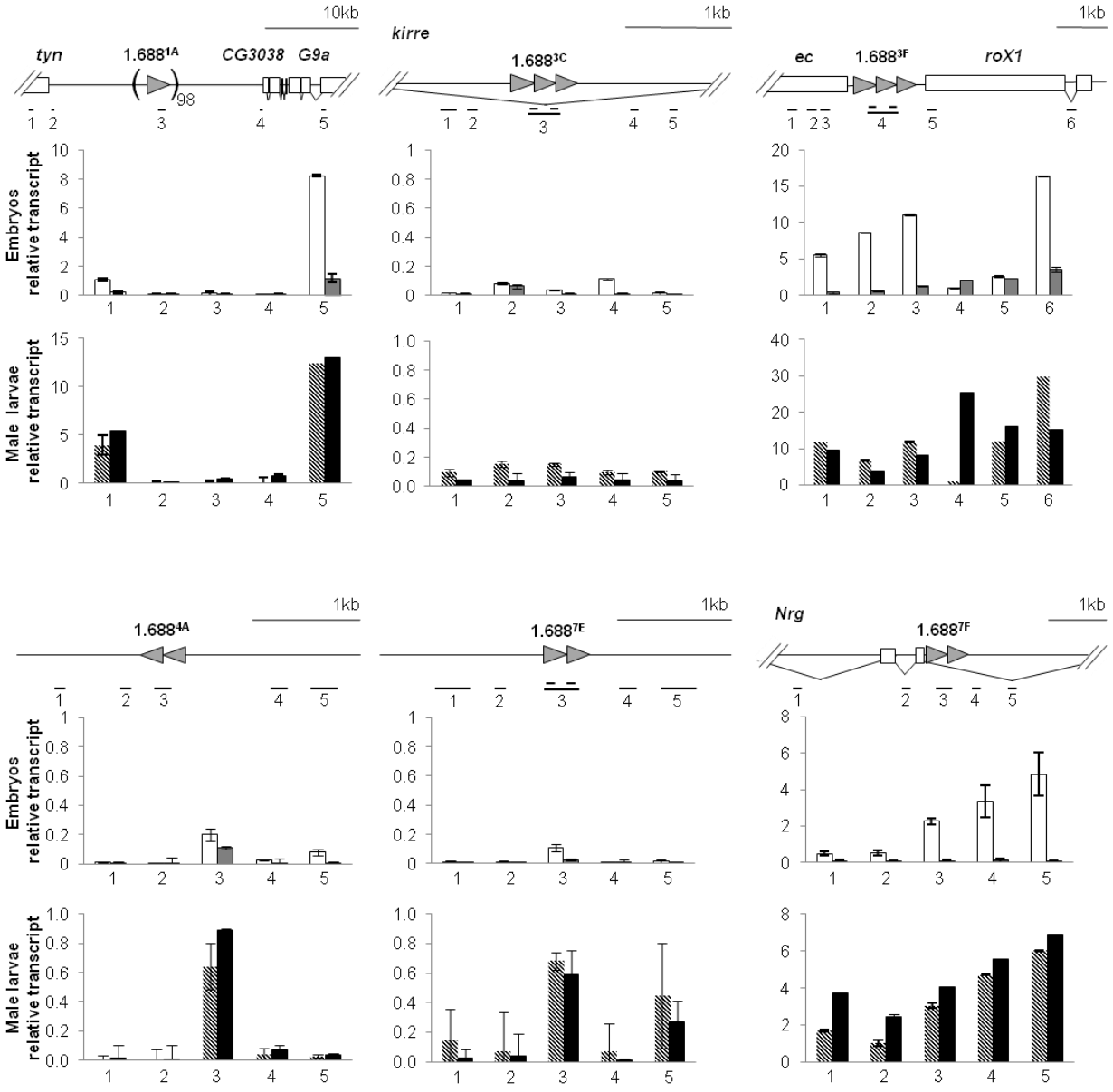


Figure 3.11. Accumulation of transcripts from 1.688^{X} repeats and surrounding regions is influenced by 1.688^{3F} siRNA. Transcript accumulation in 6-12 h embryos (white and gray bars) and male larvae (hatched and black bars) was measured by quantitative RT-PCR. White and hatched bars are controls. Gray and black bars express 1.688^{3F} siRNA. Expression is normalized to *dmn*. Standard error is derived from two biological replicates.

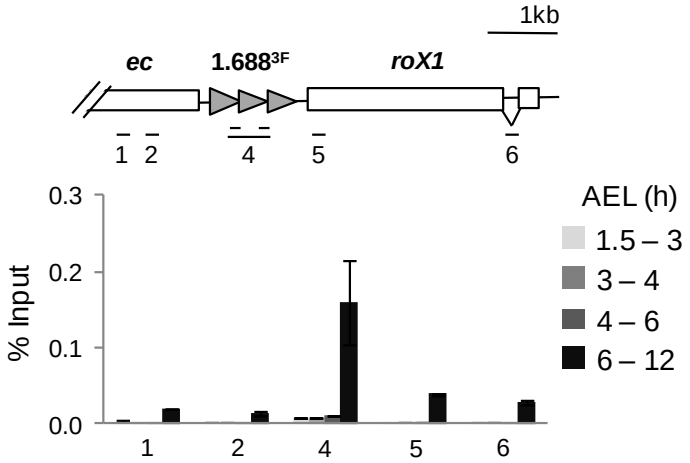


Figure 3.12. H3K9me2 deposition on 1.688^{3F} chromatin occurs 6-12 h AEL. Chromatin from staged embryos was subjected to ChIP for H3K9me2. X-localization of the MSL complex is first detected at 3 h AEL (after egg laying), but H3K9me2 enrichment is not apparent until 6-12 h AEL. Standard error is derived from two biological replicates.

H3K9me2 is enriched at regions flanking autosomal 1.688^{3F} transgenes

One challenge of studying recruiting elements on the X chromosome is that the redundancy and proximity of elements complicates interpretation. To overcome this we tested integrations of 1.688^{3F} or *roX1* on 2L (Figure 3.13 A, B, 3.14 A) (Joshi and Meller, 2017). ChIP for H3K9me2 was performed on chromatin from male 3rd instar larvae with 1.688^{3F} or *roX1* integrations on 2L (gray bars, 3.14 B, C), and in the same genotypes but with ectopic expression of 1.688^{3F} siRNA (black bars, Figure 3.14 B, C). H3K9me2 is not strongly enriched in autosomal chromatin flanking *roX1*, but is striking enriched near the 1.688^{3F} integration. Consistent with our observations in embryos, ectopic 1.688^{3F} siRNA expression elevated H3K9me2 near the 1.688^{3F} integration. This contrasts with negligible enrichment at the *roX1* transgene (Figure 3.6, Figure 3.14 B, C). For reasons that we do not understand, enrichment over the integrated 1.688^{3F} repeats was undetectable. We conclude that autosomal insertion of 1.688^{X} DNA makes nearby chromatin subject to siRNA-induced H3K9me2 deposition. Taken together, these studies support the idea that the 1.688^{X} repeats influence patterns of H3K9me2 nearby, but *roX1,* with a different class of recruiting element, does not.

To determine the influence of 1.688^{3F} and *roX1* on transcription of nearby autosomal genes, we performed quantitative RT-PCR (qRT PCR) on total RNA from 3rd instar male larvae with the 2L integrations described above, with and without ectopic 1.688^{3F} siRNA. The 1.688^{3F} and *roX1* integration site is in an intron of *haf,* one of the genes measured. We also examined *RFeSP*, *CG33128*, *Eno* (2, 58 and 114 kb from *haf,* respectively), and *CG31778* and *Rpl37A*, 2.1 and 3.5 Mb from *haf* (Figure 3.14 D). The presence of 1.688^{3F} or *roX1* integrations alone had no effect on the most distant

genes, *CG31778* and *Rpl37A*. A *roX1* integration increased expression of *haf* 2.5 fold, more than expected from full compensation. This may reflect the fact that autosomal MSL recruitment by a *roX1* transgene can overcome local, chromatin-based silencing (Kelley and Kuroda, 2003). Addition of 1.688^{3F} siRNA increased *haf* expression slightly, and similarly increases expression of *CG33128* and *Eno* (light gray bars, Figure 3.14 E).

A 1.688^{3F} insertion produced a four-fold increase in *haf*, and a slight increase in *Eno*, 114 kb distant. But upon expression of 1.688^{3F} siRNA, *haf* expression increased to 8 fold wild type levels, and *CG33128* and *Eno* both increased to ~2 fold wild type levels, consistent with full compensation. We conclude that an autosomal insertion of 1.688^{X} DNA allows relatively distant genes to increase expression, an effect that is enhanced by 1.688^{3F} siRNA.

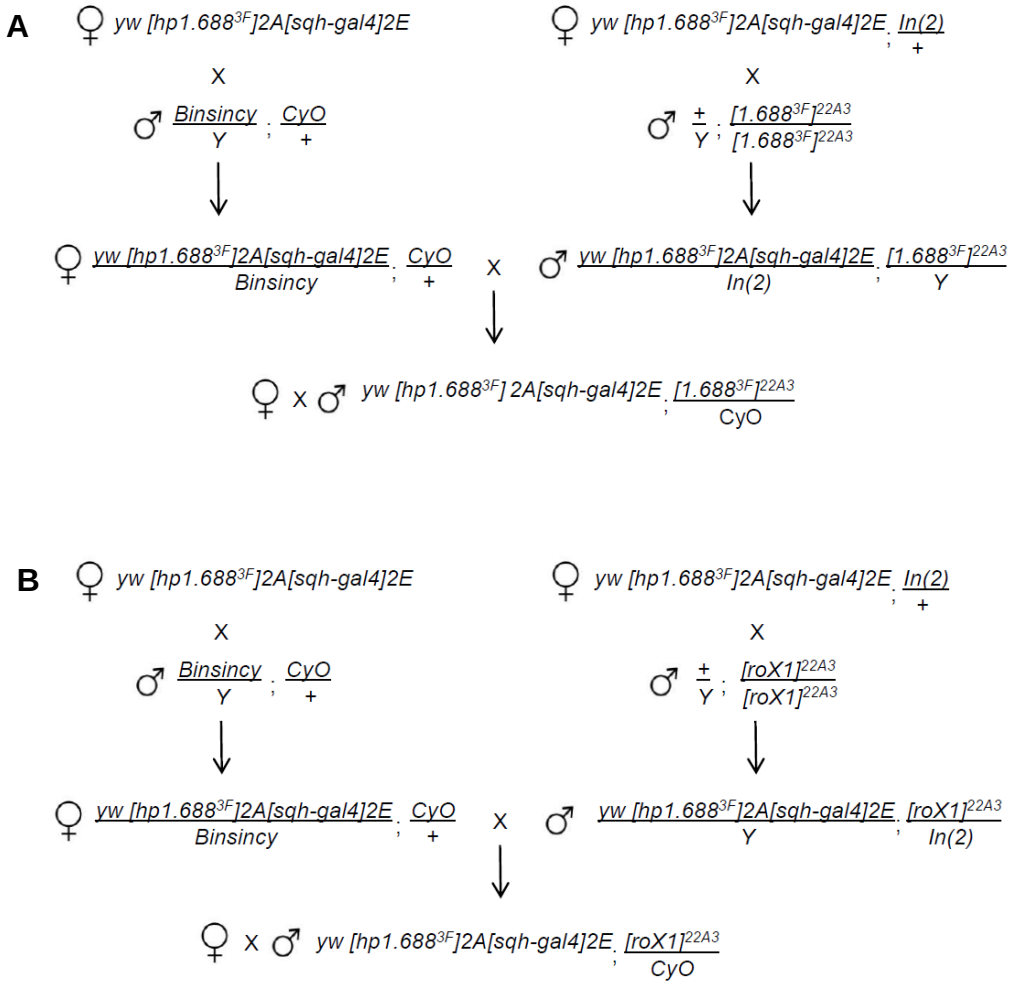
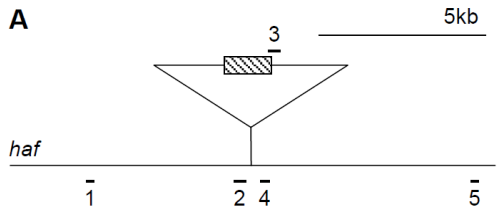
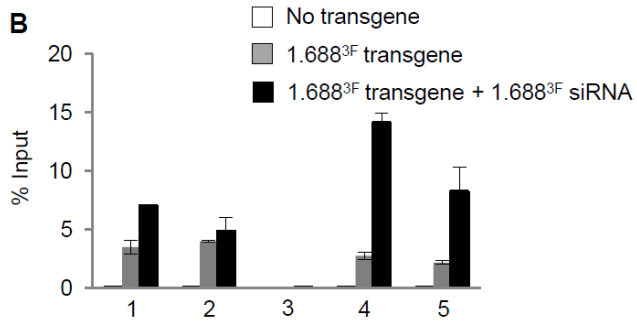
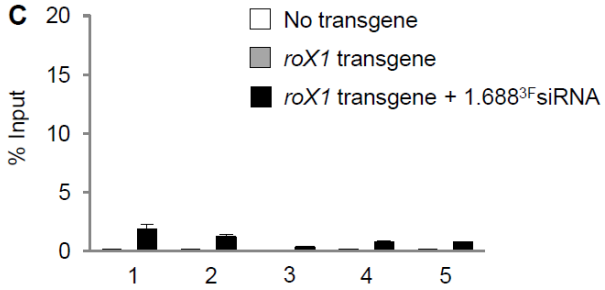
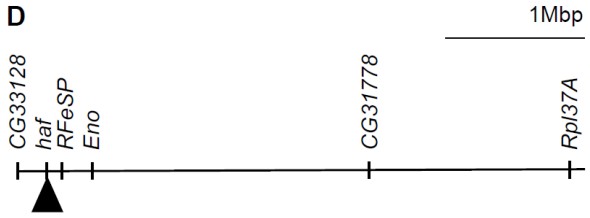


Figure 3.13. Mating scheme to generate flies that express 1.688^{3F} siRNA with 1.688^{3F} repeat and *roX1* autosomal insertion. (A-B) Virgin females carrying an X chromosome transgene that ectopically expresses 1.688^{3F} siRNA were mated with males with Binsincy X chromosome balancer and Curly, a second chromosome balancer. Virgins containing the X chromosome transgene, Binsincy, and Curly were collected. Simultaneous crosses of females carrying the same siRNA expression system on the X and the In (2L2R)*BcG* second chromosome balancer were mated to males with a 1.688^{3F} (A) or *roX1 (*B) transgene at cytological location 22A3. Male offspring from this cross with a balanced second chromosome 1.688^{3F} or *roX1* transgene were mated to sisters with balanced second chromosome transgenes. Offspring with the appropriate genotypes were collected and maintained as a stock.









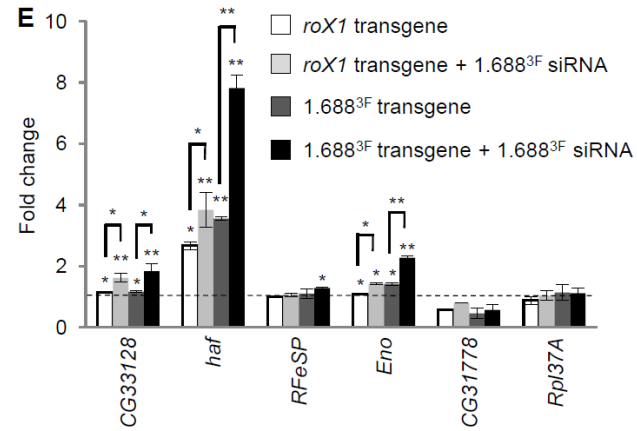


Figure 3.14. Ectopic 1.688^{3F} siRNA increases H3K9me2 flanking an autosomal 1.688^{3F} DNA insertion and elevates expression of nearby genes. (A) Amplicons flanking the landing site in a large *haf* intron at 22A3 (splicing not shown). (B) H3K9me2 enrichment surrounding the 1.688^{3F} transgene. Chromatin from wild type third instar male larvae (white), larvae with 1.688^{3F} DNA at the landing site (gray), and larvae with 1.688^{3F} DNA at the landing site and ectopic 1.688^{3F} siRNA (black) was immunoprecipitated with antibody to H3K9me2. (C) H3K9me2 enrichment surrounding a *roX1* insertion. Chromatin from wild type male third instar larvae (white), larvae with the *roX1* insertion (gray), and with the *roX1* insertion and ectopic 1.688^{3F} siRNA (black) was immunoprecipitated. Data is from two biological replicates and enrichment is normalized to input. (D) Portion of 2L showing relative location of *CG33128*, *haf, RFeSP*, *Eno*, *CG31778*, and *Rpl37A*. (E) Accumulation of transcripts in male larvae carrying *roX1* (white) or 1.688^{3F} insertions (dark gray), and in male larvae that express ectopic 1.688^{3F} siRNA and have *roX1* (light gray) or 1.688^{3F} integrations (black). Expression is normalized to *dmn*. SEM is derived from three biological replicates. Significance was determined using Student's two sample *t*-test, ≤0.05 (*), ≤0.001 (**) significance. Primers are presented in Appendix H.

Discussion

Molecularly distinct dosage compensation strategies have arisen independently in different organisms, but a shared feature is the ability to selectively recognize and alter an entire chromosome. How a regulatory system is directed to a single chromosome is poorly understood. The discovery that 1.688^{X} satellite DNA promotes recruitment of dosage compensation to nearby genes supports the idea that these repeats are important for selective recognition of X chromatin (Joshi and Meller, 2017). How the 1.688^{X} repeats accomplish this is a question of great interest. Involvement of the siRNA pathway, and siRNA from a 1.688^{X} repeat, in X recognition points to the possibility that siRNA-directed modification of chromatin around 1.688^{X} repeats plays a role in dosage compensation. The findings of the current study support this idea.

Although numerous studies point to small RNA regulation of chromatin in flies, this process is better understood in other organisms (reviewed in (Meller et al., 2015)). Small RNA directed heterochromatin formation was discovered in *S. pombe* (reviewed in (Moazed, 2009)). Heterochromatic regions are transcribed during S phase, and transcripts are processed into siRNAs that guide the Ago1-containing RITS complex to complementary, nascent transcripts (Verdel et al., 2004). In addition to several other activities, RITS recruits the H3K9 methyltransferase Clr4 (Zhang et al., 2008). We propose that a similar process is occurring at 1.688^{X} chromatin in flies. Most 1.688^{X} repeats bind Ago2, and many are transcribed. Several of the 1.688^{X} repeats that we examined are enriched for H3K9me2 deposited by Su(var)3-9, an ortholog of Clr4. Our screen identified genetic interactions between *roX1 roX2* and members of a possible RITS-like complex consisting of Ago2, Rm62 and Su(var)3-9. Finally, H3K9me2

enrichment in, and around, 1.688^{X} repeats is responsive to 1.688^{X} siRNA, and enrichment is blocked by loss of Su(var)3-9. Taken together, these findings are suggestive of a RITS-like complex modifying chromatin at 1.688^{X} repeats.

The idea that repressive H3K9me2 marks participate in a process culminating in a two-fold increase in expression is counterintuitive. However, numerous studies have found links between the compensated X chromosome of male flies and repressive marks. For example, the male X is enriched in HP1, a major constituent of heterochromatin that binds H3K9me2 (de Wit et al., 2005; Liu et al., 2005). The structure of the polytenized male X chromosome is extraordinarily sensitive to altered levels of genes that participate in heterochromatin formation or silencing, such as *HP1, Su(var)3-7* and *ISWI*. Mutations in these genes produce a general disruption of polytenization that is strikingly specific to the male X (Deuring et al., 2000; Spierer et al., 2005; Zhang et al., 2006). JIL-1, a kinase that enforces boundaries between heterochromatin and euchromatin, is enriched on the X chromosome and thought to participate in compensation (Deng et al., 2005a; Ebert et al., 2004; Jin et al., 2000; Wang et al., 2001). Upon loss of JIL-1, polytenized structure is disrupted and H3K9me2 invades euchromatic chromosome arms, but the X chromosome is most severely affected (Zhang et al., 2006). Finally, the MSL proteins themselves have an affinity for heterochromatin. In *roX1 roX2* mutant males the MSL proteins become mislocalized to ectopic autosomal sites (Meller and Rattner, 2002). For reasons that are still unclear, the most prominent of these sites are the heterochromatic 4^{th} chromosome and chromocenter (Deng and Meller, 2006; Figueiredo et al., 2014). Taken together, these observations suggest that recognition and spreading of the MSL complex could be

facilitated by repressive marks. One intriguing possibility is that 1.688^{X} repeats guide deposition of H3K9me2 and this mark, directly or indirectly, assists localization of the MSL complex.

An intriguing aspect of dosage compensation is the evolutionary convergence of mechanisms. For example, long non-coding RNA also plays a central role in X recognition in mammals, where expression of the *X inactive specific transcript* (*Xist*) RNA guides X inactivation (Lee, 2009). Furthermore, repetitive LINE-1 elements that are enriched on the mammalian X chromosome are proposed to facilitate X inactivation (Bailey et al., 2000; Lyon, 1998). Interestingly, some LINE-1 elements are transcribed during the onset of X inactivation, producing endo-siRNAs that may guide local spreading of heterochromatin into regions that are otherwise prone to escape (Chow et al., 2010). These parallels are particularly striking as the outcomes, silencing of an X chromosome in mammalian females and activation of the single X in male flies, appear unrelated. We propose that cooperation between distinct chromatin-modifying systems that rely on long and short non-coding RNAs is one strategy to selectively modulate an entire chromosome.

CHAPTER 4 SUMMARY AND PERSPECTIVES

Accurate gene regulation is a fundamental requirement for health and proper development. We have a sophisticated understanding of how individual genes are controlled, but the mechanisms that coordinately control large regions of the genome remain poorly understood. Studies of sex chromosome dosage compensation in model organisms are a valuable contribution to a comprehensive understanding of gene regulation. My research investigated a mechanism that contributes to whole chromosome recognition in flies. Significantly, this mechanism is evolutionarily conserved, suggesting a potential role in domain-wide regulation in other species.

A family of 1.688^{X} repeats is strikingly enriched on the fly X chromosome, suggesting that these might participate in X recognition. Some of these repeats have been shown to produce siRNA, and siRNA from one of these repeats acts to facilitate X recognition (Menon et al., 2014). Autosomal insertions of DNA from these repeats recruits the MSL complex to nearby genes (Joshi and Meller, 2017). The involvement of siRNA and DNA from these repeats in X recognition spurred me to ask if chromatin at the repeats is modified in an siRNA-dependent manner. I discovered that the 1.688^{X} repeats are enriched for Ago2, Ago2 interacts genetically with *roX1 roX2* mutants. Many 1.688^{X} repeats are also enriched for H3K9me2, deposited by the Ago2-interacting protein Su(var)3-9, itself displaying genetic interactions with *roX1 roX2* mutants. The H3K9me2 mark is responsive to cognate siRNA. Taken together, these observations reveal that the siRNA pathway does modulate chromatin at 1.688^{X} repeats on the X. As redundancy of repeats on the X chromosome makes it difficult to study the effect of a single repeat, we used autosomal insertions of 1.688^{X} DNA to observe chromatin

modifications and alterations in gene expression nearby. Autosomal insertion of 1.688^{X} DNA increases H3K9me2 in flanking regions and elevates expression of nearby genes. Both are further elevated by ectopic 1.688^{X} siRNA. These findings confirm the effect of 1.688^{X} repeats on flanking chromatin and expression of nearby genes. But many unanswered questions remain. The association of the repressive H3K9me2 mark with dosage compensated chromatin remains paradoxical. The mechanism by which 1.688^{X} chromatin attracts dosage compensation remains a mystery. Further investigations addressing the following questions will lead to a better understanding of mechanisms by which these repetitive elements regulate X identification in *Drosophila*.

Does an Ago2-containing effector complex localize at 1.688^{X} repeats?

We have shown that FLAG-Ago2 localizes at many 1.688^{X} repeats. In fission yeast, an Argonaute-containing RNA-induced transcriptional silencing (RITS) complex recruits a histone methyltransferase to repetitive chromatin. The RITS complex identifies these sites by base pairing between siRNA and nascent transcripts (Verdel et al., 2004). We predict that a RITS-like *Drosophila* complex is recruited to the 1.688^{X} repeats by a similar mechanism. We also demonstrated that *Su(var)3-9* places the H3K9me2 mark at 1.688^{X} repeats, but demonstrating that Su(var)3-9 protein localizes at the 1.688^{X} repeats by ChIP with anti-Su(var)3-9 would solidify this observation. My genetic screen revealed a sub-network consisting of Ago2, Rm62, Su(var)3-9, Dcr1, and Fmr1 that interacts genetically with *roX1 roX2*. It is possible that these proteins form a RITS-like complex. In support of this hypothesis I found that the Rm62 helicase is increased at 1.688^{X} repeats in embryos that ectopically express 1.688^{3F} siRNA (Appendix G). ChIP studies and co-immunoprecipitation could be performed on the

other proteins in this sub-network to examine localization at 1.688^{X} repeats and association with each other.

What is the relationship between H3K9me2 and MSL complex localization?

It is counterintuitive that the repressive H3K9me2 contributes to dosage compensation, a process that leads to hypertranscription of the male X chromosome. However, many studies have found links between repressive marks and male X chromosome. The X chromosome is enriched for HP1, a protein that binds H3K9me2 (de Wit et al., 2005; Liu et al., 2005). Jil-1, a kinase enriched on the X chromosome, maintains euchromatin and heterochromatin boundaries is also linked to dosage compensation (Deng et al., 2005a; Ebert et al., 2004; Jin et al., 2000). Upon *JIL-1* mutation, euchromatin-heterochromatin boundaries are disrupted and the X becomes particularly enriched for H3K9me2 (Deng et al., 2005a). Mutations in *Su(var)3-7,* a heterochromatin factor, disrupt the polytenized male X chromosome and reduce survival of *roX1 roX2* males (Figure 3.4, (Spierer et al., 2005)). The repressive chromatin remodeler ISWI is another protein whose mutations preferentially affect the polytenized male X (Corona et al., 2007). Although the link between repressive chromatin marks and dosage compensation is still unclear, our ability to recapitulate compensation on an autosome may allow a systematic dissection of this process.

Do the 1.688^{X} repeats affect interphase conformation of the X chromosome?

Interphase architecture of the X chromosome has been shown to have a male-specific conformation, although the details of this conformation, and how it is established, remain in dispute (Grimaud and Becker, 2009; Mendjan et al., 2006; Ramirez et al., 2015). The general consensus is that a male-specific X conformation

helps to distribute the MSL complex along the entire X chromosome. It is possible that the 1.688^{X} repeats contribute to a chromosome-specific conformation. The clade of 1.688^{X} repeats that I have studied is near-exclusive to the X, and the AT-rich sequence has high similarity to Matrix Attachment or Scaffold Attachment Regions (MAR/SAR). MAR/SAR anchors are required for the formation of chromatin loops (Heng et al., 2004). The role of 1.688^{X} repeats in affecting X conformation could be tested by Chromosome Conformation Capture (3C), a technique that reveals long-range interactions. Additionally, these repeats could be tested for association with the nuclear matrix. The 1.688^{X} repeats also have very strong similarity to a sequence favored by topoisomerase 2 (Top2), a known component of the nuclear matrix (Adachi et al., 1989; Berrios et al., 1985; Meller et al., 1995). Interestingly, mutations in *barren (barr)* a Top2 activator, enhance the male lethality of *roX1 roX2* mutants (Figure 3.4). Additionally, the proteins D1 and HP1 interact with nuclear lamins and Top2 (Blattes et al., 2006). Interestingly, D1 localizes to a group of closely related satellite repeats that makes up most of the pericentric heterochromatin of the X chromosome (Aulner et al., 2002; Blattes et al., 2006). One exciting direction for future studies is an exploration of the relationship between X conformation, the nuclear lamin and 1.688^{X} repeats on the X chromosome.

Does ectopic expression of 1.688^{X} siRNA result in global alterations in H3K9me2?

In Chapter 3 I discovered a widespread increase of H3K9me2 along the X chromosome upon expression of 1.688^{3F} siRNA, but no apparent alteration of global levels of H3K9me2 (Figure 3.10). H3K9me2 is also modestly increased at autosomal sites, sparking concern that 1.688^{X} siRNA drives a genome-wide redistribution of H3K9me2 (Figure 3.6). For example, 10 Mb of pericentric heterochromatin on the X

chromosome is composed of a closely related satellite sequence. It is possible that ectopic 1.688^{X} siRNA disrupts H3K9me2 enrichment in this region, much as it does at the euchromatic 1.688^{X} repeats. To address this we propose to use ChIP-Seq to compare the distribution of H3K9me2 in embryos from the laboratory reference strain, and embryos ectopically expressing 1.688^{X} siRNA. This will enable us to compare enrichment of H3K9me2 in heterochromatin and euchromatin genome-wide, and compare the X chromosome to the autosomes. siRNA-induced changes will be visualized using read pileups of H3K9me2 peaks in sliding window format across the genome. This will reveal whether a global redistribution of this mark occurs upon 1.688^{X} siRNA production. Heterochromatin-euchromatin boundaries will be defined by the patterns observed in the wild type embryos. Studies in a number of cell types and developmental stages reveal a remarkably distinct boundary between the euchromatic chromosome arms and pericentric heterochromatin (Riddle et al., 2011). The studies presented in Chapter 3, performed on a panel of six representative 1.688^{X} repeats, indicate that ectopic 1.688^{X} siRNA production leads to redistribution of H3K9me2 within 1.688^{X} repeats and increases H3K9me2 in surrounding regions. ChIP-seq data could be used to generate “gene models” of repeats and flanking regions to test the generality of this observation. Repeat clusters could also be sorted by the level of transcription to test the idea, put forth in Chapter 3, that H3K9me2 enrichment over 1.688^{X} repeats correlates with the level of transcription across the repeat. Although a correlation, by itself, is not proof of mechanism, a positive relationship would be suggestive of a cotranscriptional mechanism driving H3K9me2 enrichment.

APPENDIX A OPTIMIZATION OF MNASE ASSAY TO DETECT CHANGES IN CHROMATIN ACCESSIBILTY

Dosage compensation involves male-limited modification of the X-chromosome. I hypothesize that Ago2 and siRNA recruits a RITS-like complex that modifies chromatin at 1.688^{X} repeats to influence chromatin accessibility. It is possible that siRNA-dependent modification of 1.688^{X} chromatin differs in the sexes. To address this question I attempted to develop an MNase assay to determine chromatin accessibility at 1.688^{X} repeats. When implemented, this assay could also be used to explore chromatin structure at the repeats in RNAi mutants, and to determine if ectopic expression of siRNA from the 1.688^{3F} repeats influences chromatin accessibility.

Micrococcal nuclease (MNase) cleaves DNA at sites that are not occupied by nucleosomes or DNA binding proteins. Isolation of nuclei from intact animals is tricky as larvae and adults have a digestive tract full of nucleases. Isolation of embryonic nuclei is standard, but it is impractical to sex or genotype embryos. In contrast, salivary glands can be isolated from 3rd instar larvae that have been sexed and genotyped. I decided to isolate nuclei from the salivary glands of 3rd instar larvae, incubate them with different concentrations of MNase, extract DNA and perform a quantitative PCR (qPCR) with primers specific for the 1.688^{X} repeats and control regions. Accessible DNA will be cleaved at lower MNase concentrations, reducing the relative amount of template from these regions.

A hypothetical plot of template remaining after digestion with different MNase concentrations is shown in Figure A1. The y-intercept indicates the initial amount of DNA or template. A steeper initial slope indicates greater accessibility (red line, Figure A1 A). Accessibility could be inferred from the 50% destruction point where a shift to

the right would indicate less accessibility and a shift to the left would indicate more accessible chromatin. Data could also be represented on a logarithmic scale in which template remaining is indicated by RT PCR cycle number (Ct, Y axis) (Figure A1 B). The log of MNase concentration is plotted on the X-axis. This graph is expected to produce a straight line, facilitating determination of slope.

The completion of this study was intended to reveal if animal sex, Ago2 or 1.688^{X} siRNA plays a role in establishing chromatin organization at the 1.688^{X} repeats. If Ago2 modifies chromatin at the 1.688^{X} repeats then a difference in the slopes of lines following digestion of control and Ago2 mutant nuclei is anticipated (Figure A2). I may see a difference in their slopes if the male 1.688^{X} repeat differs in accessibility than the female 1.688^{X} repeat. A difference of Ct cycle in undigested samples should be observed, as the females have two copies of 1.688^{X} repeats and the males have one copy.

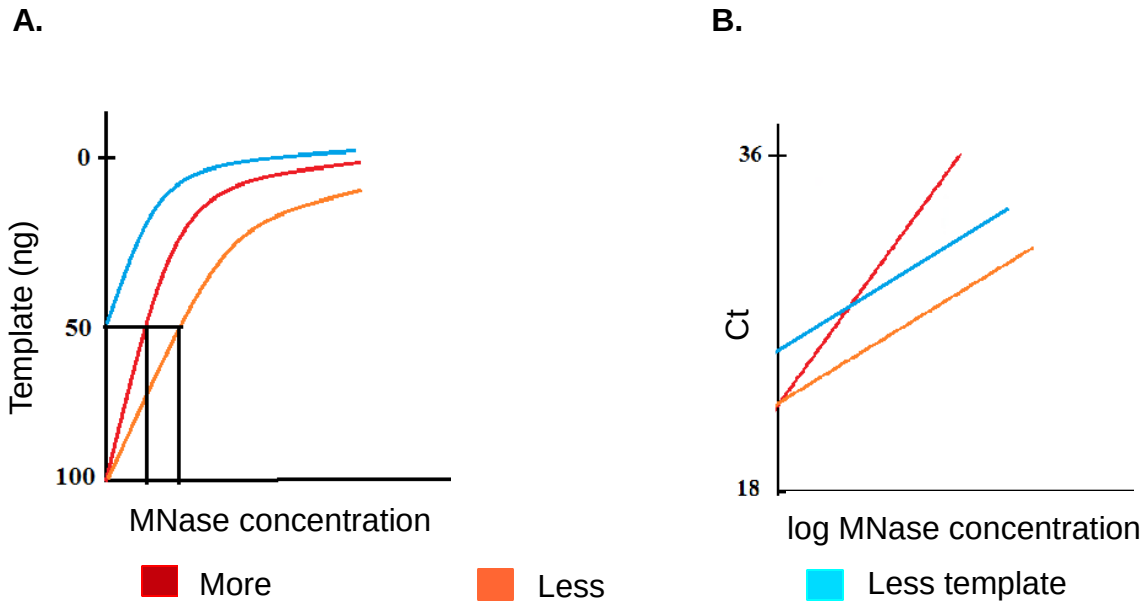


Figure A1. Comparison of conventional and logarithmic representations of MNase assay. (A) A conventional plot that shows decrease in template with the increase in MNase concentration. (B) Logarithmic plot depicting MNase concentration on the X-axis and Ct value on the Y-axis. A more accessible template is depicted by the higher slope of the red line, and the blue line depicts lower initial template concentration.

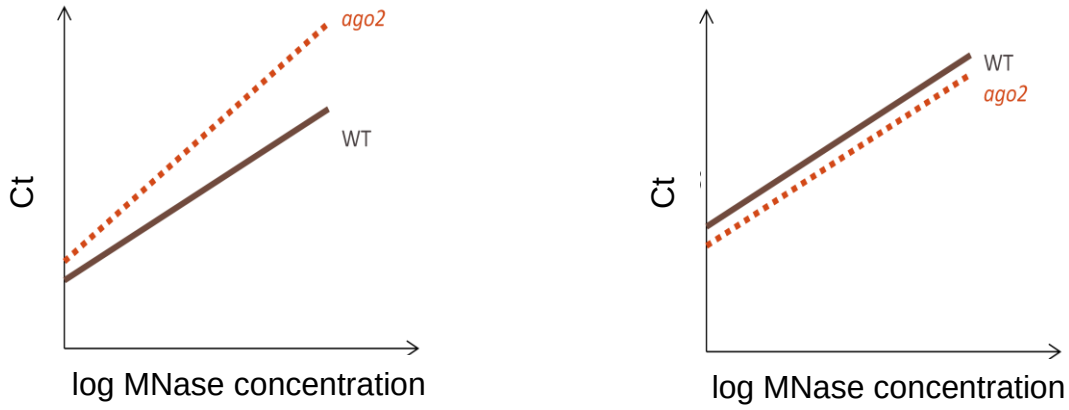


Figure A2. log-log representation of MNase assay for wild type and Ago2 mutants. The left graph depicting differences in the slopes indicate that Ago2 is necessary for manipulating chromatin structure. The right graph with parallel lines indicates that Ago2 has no role in manipulating chromatin structure.

Nuclei Isolation

A nuclei isolation protocol was adapted from (Boyd et al., 1968; Weinmann et al., 1999). The protocol was carried out at 4˚C in tubes pretreated with Nuclear Purification Buffer (NPB, 20 mM MOPS pH 7, 10 mM NaCl, 0.5 mM spermidine, 0.2 mM spermine, 3 mM $MgCl_2$) with 1 % BSA (Deal and Henikoff, 2010). Sixty pairs of salivary glands were dissected in Ringer's solution, washed and resuspended in 200 µl NPB. Eight µl 5 % Triton X-100 was added and the glands were disrupted by gentle pipetting. The suspension was passed through a nylon mesh to remove cell debris, made up to 1 ml and spun at 500 rpm for 5 min at 4˚C to concentrate nuclei. The supernatant was discarded, leaving behind about 50 µl NPB. The nuclei were washed with 50 µl of MNase Digestion buffer without $CaCl_2$ (10 mM Tris-Cl pH 7.5, 15 mM NaCl, 60 mM KCl) and collected at 6000 rpm for 10 min at 4˚C. A translucent pellet was visible after discarding the supernatant. The pellet was resuspended in 90 µl of MNase Digestion buffer with 1 mM $CaCl_2$.

Nuclei might settle or clump, producing inaccuracy. To evaluate this, 6 aliquots were mock digested, stained with DAPI and nuclei counted. The number of nuclei per 10 µl aliquot ranged from 72-99.

MNase Digestion

MNase digestion was performed as described in (Weinmann et al., 1999) with minor modifications. Five µl of MNase (NEB, Catalog #M0247S) in MNase Disgestion buffer with 1 mM $CaCl_2$ was added to 45 µl of isolated nuclei and and mixed gently. Digestion proceeded at RT for 5 min, followed by 30 min at 37˚C. The reaction was terminated by addition of 10 mM EDTA, 1 % SDS, 400 µg/ml proteinase K and tubes

were incubated at 55˚C for 2 hrs. Samples were phenol-chloroform extracted, DNA was recovered by precipitation and resuspended in 20 µl distilled water. Two µl DNA was used to template RT PCR reactions that included 10µl of SyBr Green (BioRad iTaq, #172-5101), 4 µl primers and 4 µl distilled water. Amplicons were designed to span nucleosomes. The primers are presented in Table A1.

The concentration of nuclease required and the duration of digestion was first determined. MNase concentrations between 0 and 31.25 units/ml, and digestion times of 30 and 60 min were initially tested (Figure A3). The 30 min incubation produced acceptable results (Figure A3). A linear decrease in template was observed up to 31.25 units/ml. I decided to use MNase dilutions of 10, 30, and 100 units/ml for the MNase assay (Figure A4). The autosomal gene CTCF served as a non-repetitive control that is not an expected target of siRNA-dependent chromatin modification. Amplification of CTCF, 1.688^{3F}, and 1.688^{3C} repeats produced a linear response to MNase concentration from which the slope could be calculated (Figure A5). A plasmid control consisting of serial dilutions of a plasmid containing cloned 1.688^{3F} repeat DNA was amplified at two dilutions (10^5, 10^7) chosen such that their Ct values were 22 and 28. This was developed to enable comparison between different experiments.

Although optimization of the MNase assay appeared promising, it proved impossible to replicate results in a predictable manner. As an alternative method of evaluating the status of chromatin at 1.688^{X} repeats, we switched to chromatin immunoprecipitation to determine the epigenetic marks over these regions, and explore how these marks respond to 1.688^{3F} siRNA. These studies are detailed in Chapter 3.

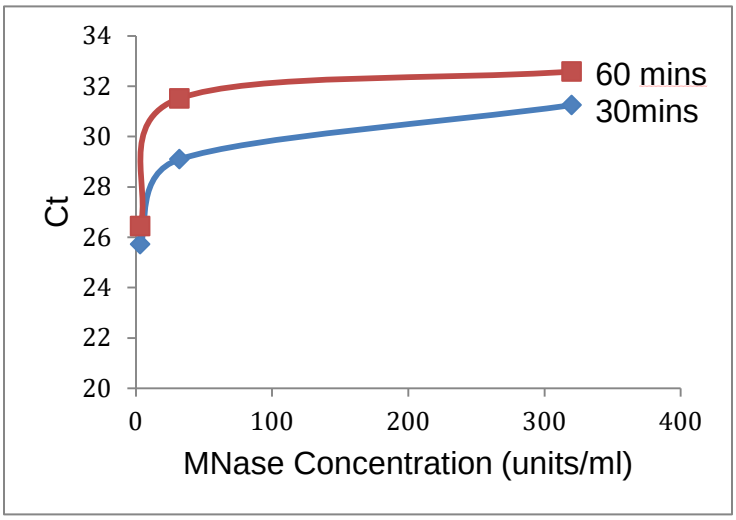


Figure A3. Optimization of MNase Digestion time. Wild type nuclei were treated with varying MNase concentrations and digestion carried out at 37˚C for 30 mins (blue) and 60 mins (red).

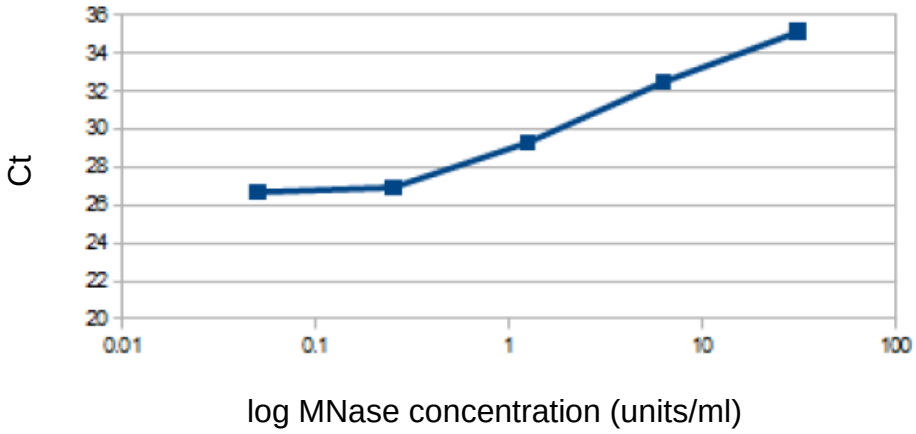


Figure A4. Determining MNase concentration range. A log-log plot of amplification of the 1.688^{3F} repeats. Log MNase concentration is on the X-axis, and Ct value on the Y-axis.

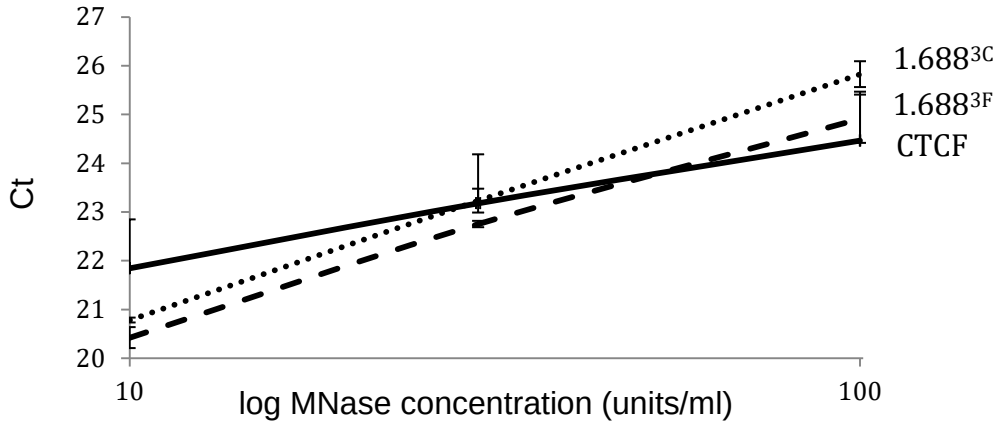


Figure A5. MNase assay in wild type flies. A log-log plot of amplification of an autosomal control – CTCF (black line), 1.688^{3F} (dashed black line), and 1.688^{3C} repeats (dotted black line).

Target	Primer	Sequence	Conc. (nM)	Efficiency
CTCF	CTCF MNase F1	GCGAGAAATCGATAAGCGC	300	107.2
	CTCF MNase R1	GTACTGACCACGGAACGTGT		
1.688^{3F}	3F MNase F1	AGCATCCACAAGAATGGGAAG	300	99.4
	3F MNase R1	TGCCAATAAACATAGCTAACTATCC		
1.688^{3C}	3C MNase F2	CGATGTTATGGCGAAAATACCGT	300	97.2
	3C MNase R2	AGAACTTAGAACGACTTTACGCA		

Table A1. List of primers used for quantitative RT-PCR studies. CTCF is used an autosomal control and primers on the 1.688^{3F} and 1.688^{3C} repeats are used as test.

APPENDIX B INVESTIGATING THE ROLE OF SU(VAR)3-9 IN X RECOGNITION

The studies described here were performed with undergraduate students Taania Girgla and Kassem Makki. This was a part of their summer research and Honors' Thesis, respectively.

Determining Gal4 driver strength

The rank strength of four Gal4 drivers was evaluated to enable variable knock down of Su(var)3-9. A stronger knockdown should more severely impact H3K9me2 enrichment, the mark deposited by the Su(var)3-9 protein. This could be used as a tool to investigate the role of Su(var)3-9 in X recognition. The genotype, stock number, and the location of different Gal4 drivers is presented in Table B1. All stocks, excepting w*; P{matα4-GAL-VP16}V2H, had an intervening sequence between the Act5 promoter and Gal4 that prevents Gal4 production. Intervening sequences, flanked by tandem Frt sites (Flip Recombinase Target), are removed by *flippase* (*Flp*), supplied by an X-linked transgene. The Flippase stock description is shown in Table B1. Excision of the intervening sequence was determined by disappearance of the y^+ marker. Gal4 strains with excised intervening sequences were mated to flies containing UAS-GFP to visually score driver strength (Table B1), and to assess the success of excision. *Flippase* efficiently removed intervening sequences. Once qualitatively scored, stocks of the four driver lines were used for Su(var)3-9 knockdown.

Genotype	BDSC Stock	Chromosome	Excised Portion	Renamed Stocks	Driver Strength
w^{118};P{AyGAL4}17b	4413	3	y	4413 ΔGAL4	++
y1 w* P{GAL4-Act5C(FRT.CD2).P}D	4779	X	CD2	4779 ΔGAL4	+++
y1 w*; P{GAL4-Act5C(FRT.CD2).P}S	4780	3	CD2	4780 ΔGAL4	++++
w*; P{matα4-GAL-VP16}V2H	7062	2	-	-	+
P{w[+mC]=ovo-FLP.R}M1A, w[*]	8727	X	-	-	-
yw	-ve Control	-	-	-	-

Table B1. **Qualitative ranking of Gal4 drivers using a GFP reporter.** Gal4 drivers were compared and ranked. The rank order of weakest (+) to strongest (++++) drivers is derived from relative GFP expression.

Knock down of Su(var)3-9

Based on the strength of the Gal4 drivers, we decided to test the weakest (w*; P{matα4-GAL-VP16}V2H) and strongest (yw; {4780ΔGal4}) drivers to create Su(var)3-9 knockdown flies. Fifty virgins with each driver were mated to male UAS-Su(var)3-9 (w^{1118}; UAS-Su(var)3-9 VDRC 101494/KK,) and TRiP-Su(var)3-9 (y^1 sc* v^1; P{TRiP.HMS00279}attP2, BDSC 33401), or to a laboratory reference strain lacking a knock down construct. No embryos obtained from matings with the strong {4780ΔGal4} driver developed into larvae. Zero-12 h and 12-24 h embryos were collected from matings using the weakest driver, P{matα4-GAL-VP16}V2H) These embryos were washed with wash solution (7 g NaCl, 3 ml Triton X-100, 997 ml water) and dechorionated using 50% bleach for 3 min. The dechorionated embryos were fixed using a 1:1 ratio of 4% paraformaldehyde and n-heptane for 25 min. The lower phase was removed, replaced with equal amount of methanol and shaken for 15 s. This step removes the vitelline membrane, allowing embryos to sink to the bottom. The supernatant and any embryos that did not sink were removed, and remaining embryos washed twice with methanol. Embryos were washed three times with PBT (PBS + 0.1% Tween 20) for 5 min each, stained with DAPI for 5 min and washed once with PBT for 15 min. PBT was removed and 10% glycerol added as the mounting medium. Embryos in different developmental stages were counted for the three matings and the percent of total embryos for each meeting determined (shown in Figure B1). The developmental profiles for the three matings did not reveal any striking differences. As similar number of females and males were used for mating, and the embryos were collected at the same time, the total number embryos obtained from each matings were also compared.

The matings with UAS-Su(var)3-9 and TRiP-Su(var)3-9 reduced the total number of embryos to 33% and 42% of that obtained for the control mating. This decrease in number of embryos could reflect strain-specific differences in female fecundity unrelated to knock down. Although decrease in number of embryos is observed in the matings with UAS-Su(var)3-9 and TRiP-Su(var)3-9 with the weakest driver, all the embryos eclosed to adult flies.

Next, I performed qRT-PCR to quantify Su(var)3-9 transcript. Su(var)3-9 has 5 exons, the first two unique to Su(var)3-9 and the last 3 shared with eIF2γ. I designed primers spanning the 1st and the 2nd exon (Table B2). Relative transcript in 3rd instar larvae from matings between w*; P{matα4-GAL-VP16}V2H and UAS-Su(var)3-9, and w*; P{matα4-GAL-VP16}V2H and TRiP-Su(var)3-9 showed decrease in Su(var)3-9 transcripts (Figure B2).

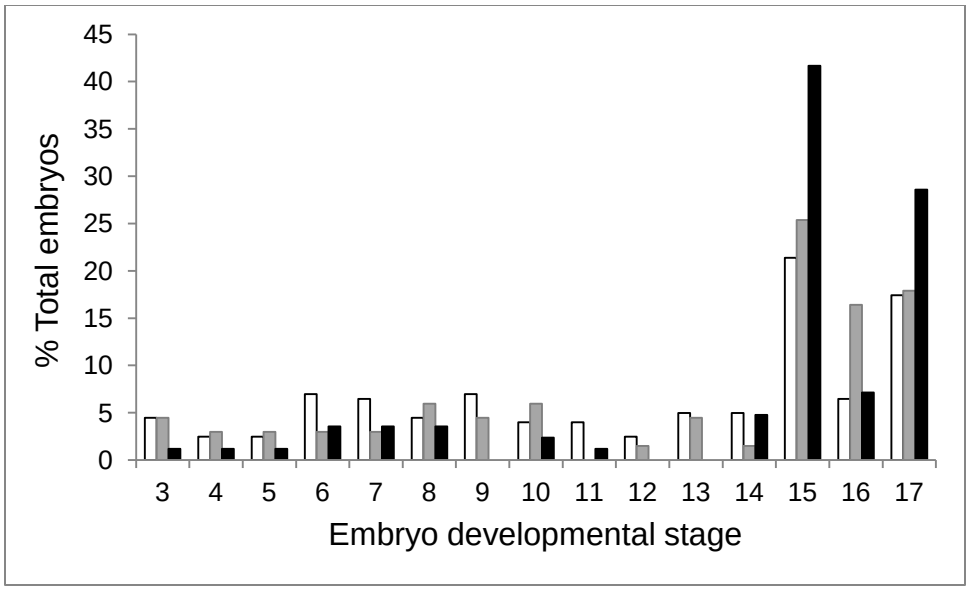


Figure B1. Embryo count from different developmental stages of Su(var)3-9 knockdown embryos. Percent total embryos obtained from matings between laboratory reference strain (yw) and the weakest driver (w*; P{matα4-GAL-VP16}V2H) are shown in white. Gray bars show percent total embryos from matings between UAS-Su(var)3-9 and the weakest driver, and black bars show percent total embryos from matings between TRiP-Su(var)3-9 and the weakest driver.

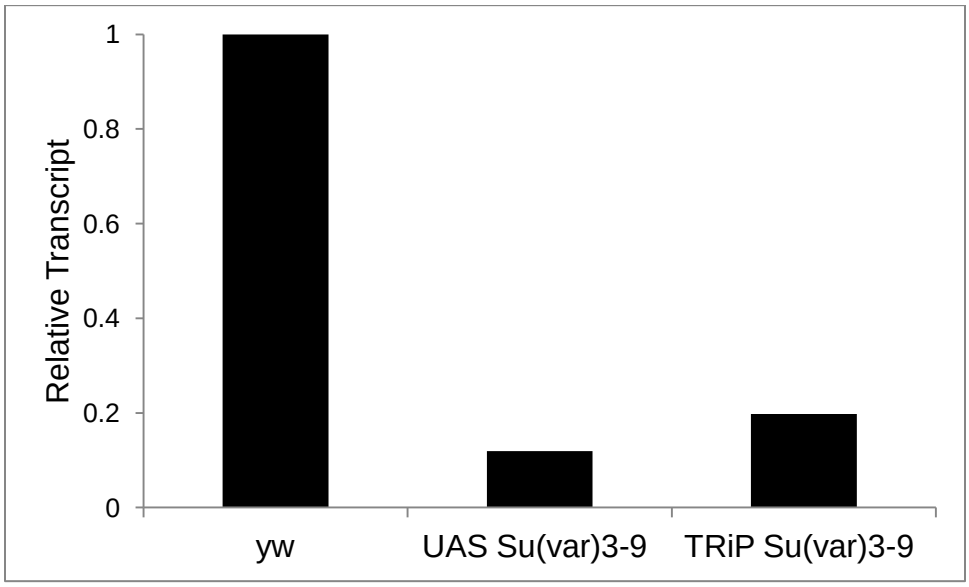


Figure B2. Accumulation of Su(var)3-9 transcripts is decreased in flies that are knockdown for Su(var)3-9. Transcript accumulation in 3rd instar larvae was measured by quantitative RT-PCR.

Generating Su(var)3-9 knockdown flies.

Although we were successful in achieving Su(var)3-9 knockdowns, fly stocks used to achieve Su(var)3-9 knockdown also target an essential overlapping gene, eIF2γ. Because of this, knockdown with a strong driver caused lethality. To limit knock down to Su(var)3-9, I worked with Kassem Makki, an undergraduate honors student. His study focused on generating an RNAi knockdown construct that targets an exon exclusive to Su(var)3-9. This would allow elimination of Su(var)3-9, without disrupting eIF2γ.

To accomplish this, an exon present only in Su(var)3-9 was cloned into a pWIZ vector. pWIZ is designed to facilitate generation of transgenic flies for RNAi knockdown (Bao and Cagan, 2006). The pWIZ vector allows generation of dsRNA in flies, which is processed into siRNA. This vector will be injected in fly embryos to create strains that knockdown Su(var)3-9 when combined with a Gal4 driver.

We amplified a 450 bp segment from the 2nd exon of Su(var)3-9 with primers that would create AvrII restriction enzyme sites. The PCR product was purified and cloned in a pCR4 Topo cloning vector, enabling us to sequence and amplify the insert before cloning into pWIZ. The pWIZ vector was digested with AvrII and ligated to an AvrII-digested insert. Colony PCR was performed using a reverse primer in the insert and a forward primer in the hsp70 promoter present in pWIZ) (Table B2, Figure B4). A clone with insert was identified and used for the second cloning step at Nhe1.

Cloning in NheI site proved to be challenging. After many failed attempts we decided to use an alternative approach to show role of Su(var)3-9 in depositing H3K9me2 at the repeats (Chapter 3, Figure 3.9).

Primer name	Sequence	Working Conc. (nM)	Efficiency
Su(var)3-9 F2	CCACGGTGGTCAAAGCCATA	300	92.5
Su(var)3-9 R2	CTGCTGCTTGGAGGTCAAAAG	500	
Hsp 70 promoter F	GAGAGAATTCCCCCTAGAATCCCAAAA C		
Su(var)3-9 ex AvrII FP1	ATTCCATGGGTAGATAGACGCACCACC CG		
Su(var)3-9 ex AvrII RP1	TAACCATGGATGCGCTTCTCGAACAAT GC		

Table B2. Primer table. Primer Su(var)3-9 F2, R2 were used for qRT-PCR. Hsp70 promoter, Su(var)3-9 ex AvrII primer pairs were used for cloning in pWIZ.

APPENDIX C DETERMINATION OF GENOME WIDE ALTERATIONS IN H3K9ME2 LEVEL IN FLIES EXPRESSING 1.688^{3F} siRNA

The major limiting step towards carrying out ChIP-Seq is that the protocol we were using did not yield enough double stranded DNA to make ChIP-seq libraries. Ten ng of dsDNA is required for library generation. To get ChIP-seq compliant DNA, I standardized the steps of chromatin generation, resulting in the protocol detailed below.

PROTOCOL FOR CHROMATIN PREPARATION

Overnight embryos (0.35 g) were collected, dechrorionated with 50 % bleach for 3 min, and crosslinked for 20 min in 50 mM HEPES, 1 mM EDTA, 0.5 mM EGTA, 100 mM NACl, 1 % formaldehyde with heptane and vigorous shaking. Crosslinking was stopped by adding 125 mM glycine, 0.01 % Triton X-100 in 1 X PBS for 30 min. Crosslinked embryos were washed once in PBS and collected at 2500 rpm for 5 min at 4˚C. Supernatant was discarded and embryos were resuspended in 15 ml Buffer B (Table C1) and 1 Complete^{TM} protease inhibitor tablet (Roche, #11697498001), followed by incubation on ice for 15 min. Embryos were homogenized in a precooled Dounce, 10 strokes with pestle A followed by 15 strokes with pestle B. The homogenized mixture was centrifuged at 170 g for 10 min at 4˚C and supernatant discarded. The pellet was washed twice with MNase Digest Buffer (Table C1), resuspended in 200 µl of MNase Digest Buffer and transferred to a 2 ml tube. DNA was digested with 40 Units of MNase (2.67 µl of 15 units/ µl, Wortington Stock #LS004797) 37˚C for 30 min. Digestion was stopped by adding 12 µl of 0.5 M EDTA, and the volume made up to 1 ml using Ten140 buffer (Table C1). The sample was sonicated by a Fischer Scientific Model FB505 on ice at 35% amplitude, total time of 2.5 min, 30 sec on, 59 sec off, producing shearing to

200-600 bp (Figure C1). Sonicated chromatin was centrifuged at 12,000 rpm for 15 min at 4˚C and 200 µl aliquots stored at -80˚C.

Chromatin Immunoprecipitation

Two hundred µl of chromatin is used for each chromatin immunoprecipitation, which was carried out as detailed in Chapter 3 with the exception of reverse crosslinking. To reverse crosslinks of eluted samples, 4 µl RNase A (10 µg/µl, invitrogen PureLink™ Stock #12091-021) was added and samples were incubated at 37˚C for 15 min. Twenty µl of 5 M NaCl was added and samples were incubated at 65˚C for 6 hr. Two µl of proteinase K (10 µg/µl, Thermo Scientific Stock #26160), 10 µl 0.5 M EDTA, and 10 µl 1 M Tris (pH 7.6) were added and samples incubated at 45˚C for 1 hr. Extraction with an equal volume of phenol:chloroform:isoamylalcohol (25:24:1 v/v) was performed, and the aqueous phase transferred to another 0.5 ml tube. Ten µl glycogen, 1/10 volume of 3 M sodium acetate (pH 5.2) and 2.5 volume of chilled 100% ethanol were added. DNA was precipitated at -20˚C for at least one hour and pelleted at 14,000 rpm for 15 min at 4˚C. The pellet was resuspended in 50 µl distilled water and stored at -20˚C.

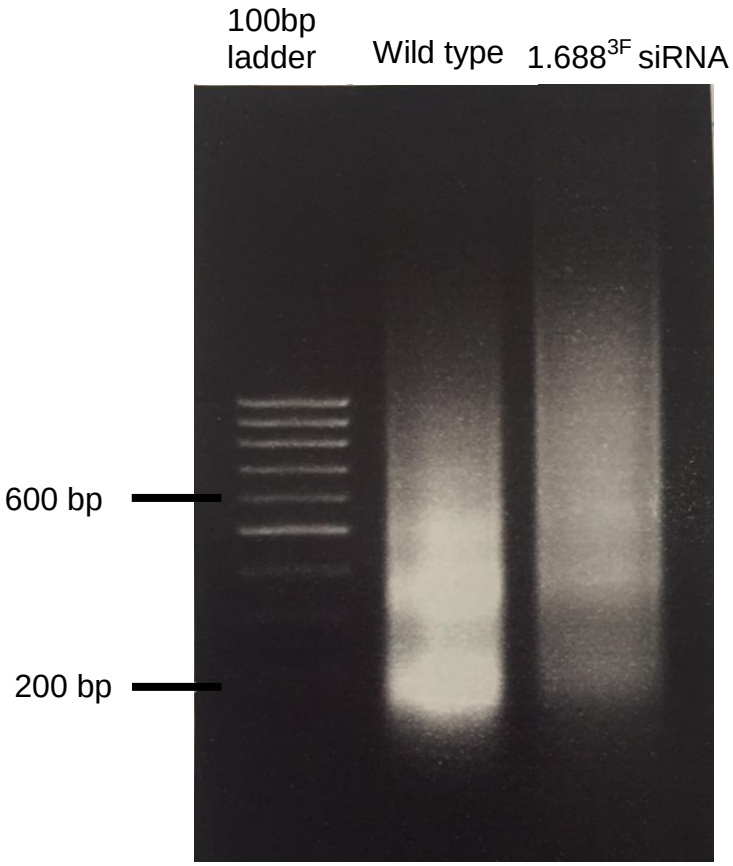


Figure C1. Reverse crosslinked chromatin. Chromatin was prepared from 6-12 hr embryos obtained from wild type embryos (left) and embryos expressing 1.688^{3F} siRNA (right) using the revised protocol. Shearing between 200-600bp was obtained.

Re-suspension Buffer B

Components	Stock	50 ml
10mM Tris (pH 8)	1M	500 µl
10mM KCl	1M	500 µl
3mM $CaCl_2$	100mM	1.5ml
0.34M Sucrose	1 M	5.81g
1mM DTT	1M	50 µl
0.1% Triton X100		50 µl
0.2mM EGTA	0.5M	20 µl
ddH_2O		Add to 50 ml

MNase Digest Buffer

Components	Stock	50 ml
15mM Tris (pH 8)	1M	750 µl
60mM KCl	1M	3ml
15mM NaCl	5M	150 µl
1mM DTT	1M	50 µl
0.25M Sucrose		4.27gm
1mM $CaCl_2$	100mM	500 µl
ddH_2O		Add to 50 ml

Ten140 buffer

Components	Stock	50ml
140mM NaCl	5 M	1.4 ml
10mM Tris (pH 7.6)	1 M	500 µl
2mM EDTA	0.5 M	200 µl
ddH_2O		Add to 50 ml

Table C1. Buffers used for chromatin preparation.

APPENDIX D DETERMINING REPEAT LENGTH AND COPY NUMBER IN THE LABORATORY REFERENCE STRAIN

A clade of 1.688^{X} repeats is dramatically enriched on the X chromosome and involved in X recognition. These are present in short clusters of ~359 bp tandem repeats that are distributed throughout euchromatin. While the position of repeat clusters is stable, there is variation in the number of repeats within each cluster, and this can be detected by comparing different strains of *D. melanogaster*. To support the studies described in Chapter 3, I wished to determine the number of repeats in our laboratory reference (*yw*) strain at several cytological positions. This was done by PCR of DNA using primers flanking each repeats (Table D1). The length of the amplicon attributable to repeats was divided it by 359 to obtain the copy number. Prior studies have shown that the copy number of 1.688^{3F} repeats in the laboratory reference strain is 3.5. The lengths of 1.688^{4A}, 1.688^{3C}, 1.688^{7E}, and 1.688^{7F} repeats are 850 bp, 1200 bp, 900 bp, and 500 bp, respectively, bringing their copy numbers to 2.5, 3.5, 2.5, and 1.5 (Figure D1). As the 1.688^{1A} repeats have a copy number of ~100, it was not possible to use this method to determine 1.688^{1A} repeat length or copy number in the laboratory reference strain.

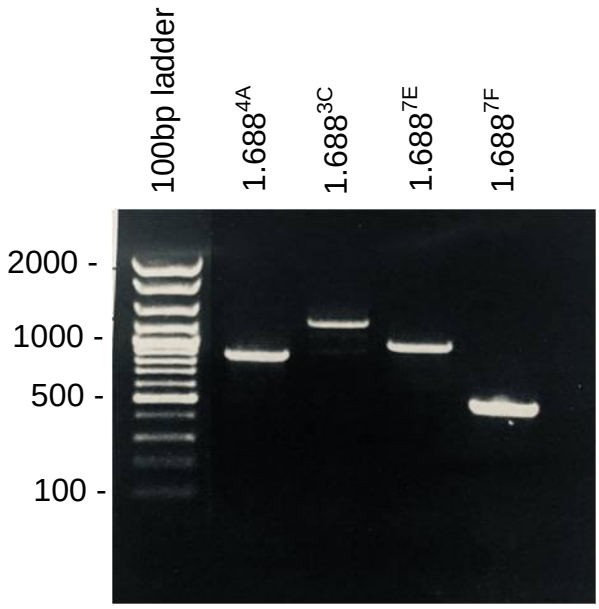


Figure D1. Determination of tandem repeat size in laboratory reference strain. Amplicons produced from laboratory reference strain are used to determine the number of tandem repeats in 1.688^{4A}, 1.688^{3C}, 1.688^{7E}, and 1.688^{7F} in the laboratory *yw* reference strain.

Primer Name	Sequence	Amplicon Size	Approx. repeat copy number
4A Across F1	AGTGCGAGGTACACCGAAAG	890	2.5
4A Across R1	ACCGAACAACATTCGGGCAT		
3C Across F1	GACATACATCGTTGAGTTCGCA	1250	3.5
3C Across R1	TGCCAAGCTTATAACTACTGCT		
7E across F1	ACGAACCCTATAACTTTTTAACGCA	910	2.5
7E across R1	TGATACCAATCAAGTGGTCTAATGA		
7F across F1	AGTCCTTCCAAAAGTGATAGCG	550	1.5
7F across R1	CGCTAGAAAGGATCACTTCTTTTCA		

Table D1. Primer table to determine repeat length. Primer names and sequences used are listed. Primers were made flanking each repeats. Amplicon size deduced from Figure D1 is shown. Approximate repeat copy number is determined by dividing amplicon size by 359bp.

APPENDIX E INCLUSION CRITERIA FOR GENETIC SCREEN OF AGO2 INTERACTORS

Physical interactions between proteins are necessary for cellular processes. Numerous studies have experimentally deduced many protein-protein interactions. Significant efforts have been made to develop databases that curate and archive these experimentally discovered protein-protein interactions. Biological General Repository for Interaction Datasets (BioGRID) (Stark et al., 2006) is one such curated database that is widely used. Different databases use different rules and systems to compile data. In addition, networks that integrate major databases have been developed. This allows the user to retrieve interaction data using a simple user interface. One example of this, esyN (Bean et al., 2014), has been adopted by Flybase (http://flybase.org). I used BioGRID and esyN to generate a list of Ago2 interactors (Table E1). These sources provide information on interactions from researchers with different methods and interests, and they may or may not be relevant for our study. A manual curation of the list of Ago2 interactors was required.

Four criteria were used to rank the Ago2 interactors identified in BioGRID (Stark et al., 2006) and esyN (Bean et al., 2014). The means of detecting interaction was scored 0 if high throughput and 1 if low throughput or validated. Additional criteria are known roles in RNA interference, chromatin modification and association with chromatin. Proteins received a score of 1 for each criterion satisfied and 0 if unsatisfied. Values were added and presented as the Curation Score in Table E1. Most named proteins with a curation score of 2 or more were tested for genetic interaction with *roX1 roX2*. Phosphatidylinositol glycan anchor biosynthesis class S (PIG-S, curation score 1) was also tested. No Ago2-interactors with a score of 1 or 2

displayed genetic interactions with *roX1 roX2* mutants, but several proteins with a score of 3 or 4 enhanced the lethality of *roX1 roX2* males, suggesting a role in dosage compensation (Table E1, Figure 3.4).

Protein	Symbol	EsyN ID	Database	Curation score	*roX1 roX2* interaction
Dicer-1	Dcr-1	FBgn0039016	BioGRID	4	Yes
Dicer-2	Dcr-2	FBgn0034246	esyN	4	Yes
Elongator complex protein 1	Elp1	FBgn0037926	esyN	4	Yes
Fragile X mental retardation	Fmr1	FBgn0028734	BioGRID	4	Yes
r2d2	r2d2	FBgn0031951	esyN	4	No
Rm62	Rm62	FBgn0003261	BioGRID	4	Yes
Small ribonucleoprotein particle protein 1	SmD1	FBgn0261933	esyN	4	No
TBP-associated factor 11	Taf11	FBgn0011291	esyN	4	No
vasa intronic gene	vig	FBgn0024183	esyN	4	Yes
barren	barr	FBgn0014127	esyN	3	Yes
belle	bel	FBgn0263231	esyN	3	No
CTCF	CTCF	FBgn0035769	esyN	3	No (previously tested)
smaug	smg	FBgn0016070	esyN	3	Yes
forkhead box, sub-group O	foxo	FBgn0038197	esyN	2	No
p53	p53	FBgn0039044	BioGRID	2	No
I-kappaB kinase β	IKKβ	FBgn0024222	BioGRID	2	-
CG9302	CG9302	FBgn0032514	BioGRID	2	-
Phosphatidylinositol glycan anchor biosynthesis class S	PIG-S	FBgn0265190	BioGRID	1	No
Centrosomal protein 190kD	Cp190	FBgn0000283	esyN	1	-
DNA fragmentation factor-related protein 2	Drep2	FBgn0028408	esyN	1	-
gigas	gig	FBgn0005198	esyN	1	-
Negative elongation factor E	Nelf-E	FBgn0017430	esyN	1	-
Phosphatase and tensin homolog	Pten	FBgn0026379	esyN	1	-
Phosphoinositide-dependent kinase 1	Pdk1	FBgn0020386	esyN	1	-
Proteasome α7 subunit	Prosα7	FBgn0023175	esyN	1	-
pumilio	pum	FBgn0003165	esyN	1	-
Purine-rich binding	Pur-α	FBgn0022361	esyN	1	-

Protein	Symbol	EsyN ID	Database	Curation score	*roX1 roX2* interaction
protein-α					
rapamycin-insensitive companion of Tor	rictor	FBgn0031006	esyN	1	-
Ras homolog enriched in brain	Rheb	FBgn0041191	esyN	1	-
Ribosomal protein S6 kinase II	S6kII	FBgn0262866	esyN	1	-
RNA polymerase II 215kD subunit	RpII215	FBgn0003277	esyN	1	-
Tsc1	Tsc1	FBgn0026317	esyN	1	-
Tudor staphylococcal nuclease	Tudor-SN	FBgn0035121	esyN	1	-
Ubiquitin-conjugating enzyme variant 1A	Uev1A	FBgn0035601	esyN	1	-

Table E1. Ago2-interactors ranked by manual curation. Ago2 interactors were identified from publically available data bases, manually curated by criteria described in the text and ranked by curation score. The right hand column indicates whether a genetic interaction with a *roX1 roX2* mutant chromosome has been detected. A (-) indicates that the gene has not been tested for interaction with *roX1 roX2*.

APPENDIX F DETERMINING EFFECT OF AGO2 ON H3K9ME2 ENRICHMENT AT 1.688^{X} REPEATS

Males with the partial loss of function *roX1*^{ex40}*roX2Δ* chromosome have normal survival, but synthetic lethality is observed if both copies of *ago2* are mutated. This is accompanied by reduced X-localization of MSL complex (Menon and Meller, 2012). In Chapter 3 we observed Ago2 localization at many repeats. This prompted us to ask, what happens to H3K9me2 at the 1.688^{X} repeats upon loss of Ago2? To test this, I collected male third instar larvae from *Ago2*$^{-/-}$ and wild type flies and performed chromatin immunoprecipitation using anti-H3K9me2 (ab1220). The protocol used is described in Chapter 3. This was followed by quantitative PCR with primers listed in Appendix H.

Contrary to expectation, these studies reveal that H3K9me2 enrichment increases over 1.688^{X} repeats, as well as in some flanking regions (Figure F1). Although we do not understand why this is happening, repeat regions are typically silenced by H3K9me2. I hypothesize that in the absence of Ago2, an alternative silencing mechanism takes over. Future studies might be aimed at identifying the methyltransferase responsible for placing this mark in *Ago2*$^{-/-}$ larvae.

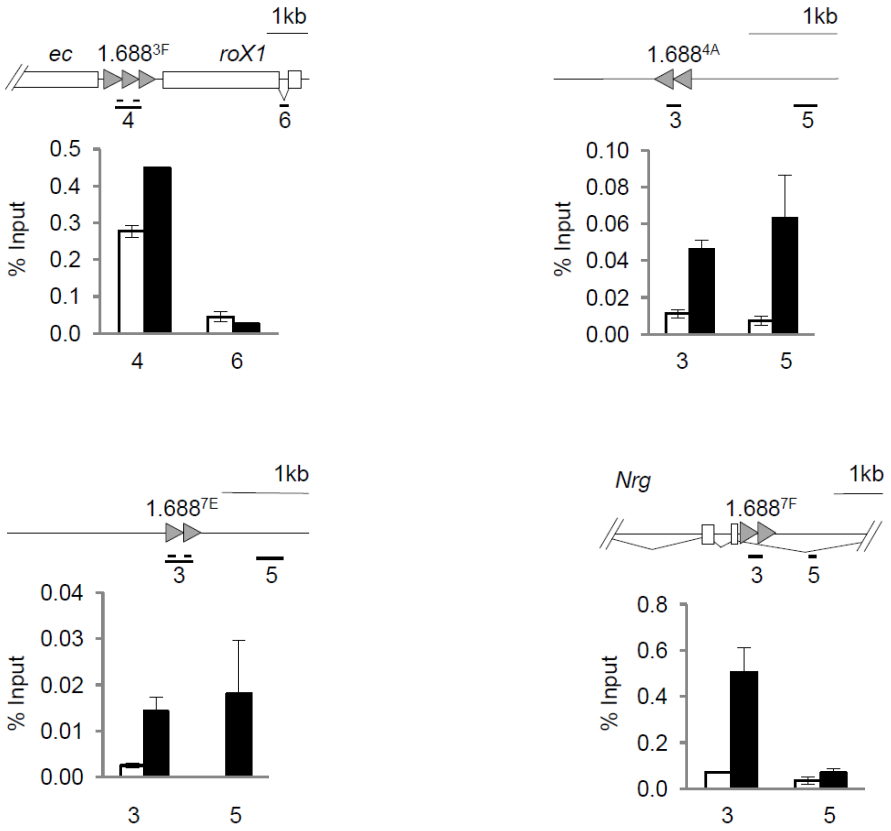


Figure F1. Loss of Ago2 increases H3K9me2 at 1.688^{X} repeats. Chromatin from wild type male larvae (white), and *Ago2*$^{414/414}$ male larvae (black) was immunoprecipitated with antibody to H3K9me2. Enrichment normalized to input is shown. Standard error is derived from two biological replicates.

APPENDIX G LOCALIZATION OF RM62 AT 1.688^{X} REPEATS

The genetic screen detailed in Chapter 3 revealed a sub-network composed of Ago2, Rm62, Su(var)3-9, Dcr1, and Fmr1 that interacts genetically with *roX1 roX2* mutants. Ago2 localization at the 1.688^{X} repeats, and deposition of H3K9me2 at these repeats by Su(var)3-9, prompted us to ask whether Rm62 also localizes to 1.688^{X} chromatin. We performed ChIP on 6-12 hr wild type embryos and embryos expressing 1.688^{3F} siRNA using 8 µl anti-Rm62 (Abcam, ab52809) antibody. The protocol used is as described in Chapter 3. This was followed by quantitative PCR using the primers in Appendix H. I found that Rm62 is increased at and around 1.688^{X} repeats by increased 1.688^{3F} siRNA.

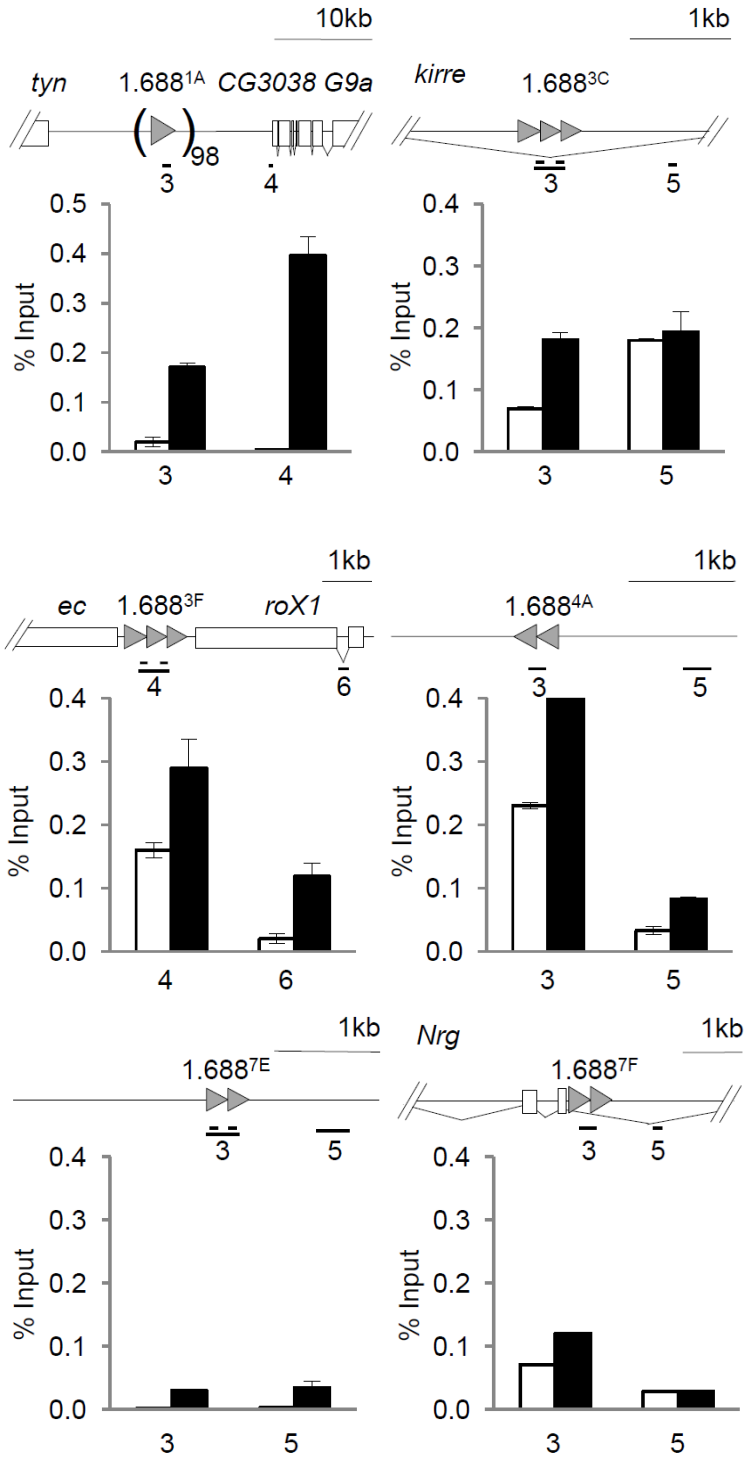


Figure G1. Elevated 1.688^{3F} siRNA increases Rm62 localization over and around 1.688^{X} repeats. Chromatin from wild type embryos (white) and embryos ectopically producing 1.688^{3F} siRNA (black) was immunoprecipitated with antibody to Rm62. Enrichment over input was determined by quantitative PCR (qPCR). The standard error of two biological replicates is shown. Primers used for analysis are presented in Appendix H.

APPENDIX H PRIMER LIST

A. Primers used for ChIP-qPCR analysis.

ChIP for FLAG-Ago2

Gene name	Primer name	Sequence	Working Conc. (nM)	Efficiency
Hsp70	hsp70 - 200F	TGCCAG AAAGAAAACTCGAGAAA	300	95.6
	hsp70 - 108R	GACAGAGTGAGAGAGCAATAGTACAGAGA		
dmn	Dmn P2 left	AGATTGAGCAGAAGCAGGGA	300	99.8
	Dmn P2 right	CAGCAGCTCCTTGTTGTTCA		

ChIP at 1.688^{X} repeats

Amplicon No.	Primer name	Sequence	Working Conc. (nM)	Efficiency
ChIP at 1.688^{1A1}				
1	1A F2 3' Tyn	GGAATAAGCTGCGAGCCCGTAC	150	91.3
	1A R2 3' Tyn	AATGTGGTCTCGTGTGAGTACGTAA		
2	1A1 L ND F1	AGTGCTCTGTGTGCATTGGT	500	103.2
	1A1 L ND R1	GTGGCGAAGCCAGTTTTCAG		
3	1A1 ND F1	TCCGATTTTTGGCAAT	500	98.2
	1A1 ND R1	AAGCGTAAATGAAGAC		
4	1A1 R ND F1	TGTCTTAGCCTTTAGAACTAAGTGT	300	103
	1A1 R ND R1	CGACAAAACGCGGAATGTCTT		
5	1A F4 G9a	GACACGCCCACTTCAGTTACTGATG	300	97.4
	1A R4 G9a	CGGGTCTTATTTTCCTGGCTCG		
ChIP at 1.688^{3C}				
1	3CL F2	TTTTTAGCTATGCCCCGCGA	150	93.3
	3CL R2	GGCAAGCGGAAACACTGAAG		
2	3CL F1	CTGGCGTGAATGTAGCTCGTAA	300	98.2
	3CL R1	GCTCCGTTTCTCTGCCGTATT		
3	3C F1	CCAATCCAACTGTAACCCCGA	500	94.8
	3C R1	TTTGTAAGGGGTAACATCATGAAAA		
4	3CR F2	TCAAGGATGCTGCGGTTTTG	300	96.8
	3CR R2	CACTCCAGCATGCAGGTTAAT		
5	3CR F3	CCAACTTGTGTGGCTAAGCTC	300	94.2
	3CR R3	GCCGCTTTTAGTCGGATTTCA		

Amplicon No.	Primer name	Sequence	Working Conc. (nM)	Efficiency
ChIP at 1.688[3F]				
1	Ec F2	GCCAGCTAACAGGCGAATTG	300	95.1
	Ec R2	GTCACAGCGGGAACTTCTCA		
2	Ec F3	AGTGTTGCGACTTCAGAGCA	300	97.9
	Ec R3	ATGTTGCTGGGCATTGGGTA		
3	Ec F5	TGAACCAGCGCAGATGATGA	300	99.1
	Ec R5	TCCTTGGCGGCTCCTTATTG		
4	RNR1 F1	TATTTACAAACGGGGTTATCTCTATAAGG	500	100.5
	RNR1 R3	CGTAACAAAATTTCCTATCGACCT	300	
5	3F ChIP F3	TCGGCTCAGGCGTATAACGA	300	102.2
	3F ChIP R3	TGAAATGAACACAGCCAAAGCA		
6	roX1 ChIP F2	TGCCGCCAAAGACTGATGAT	300	101
	roX1 ChIP R2	CCTTGACGAGTCCGGACAAT		
ChIP at 1.688[4A]				
1	4A L ND F3	GCCATTCCCCTCCCCAGTTA	300	92.5
	4A L ND R3	GCGATTGCTGTGCCATTTCA		
2	4AL ND F'2	CCGCCTCTGTCGTACTTTCA	150	90
	4AL ND R'2	TCATTTCCTTCGGCTTGGCT		
3	4A ND F1	AAGTCTCGTAGGACGCAGGA	150	96.9
	4A ND R1	GTACCTCGCACTTGCTGACT		
4	4A R ND F2	CATTTGTCTGCTGCGTGAGC	300	101.6
	4A R ND R2	TGCTGCGTCTTGACTTTCGT		
5	4A R ND F4	CATTGAAGCGGTTGCGGATT	150	98.7
	4A R ND R4	TGGGGTTATTTTCGGAGGGC		
ChIP at 1.688[7E]				
1	7E L ND F4	AGATGTGGTCAAACACTGCG	300	104.3
	7E L ND R4	AAACCGAAACCGAGAACCAGA		
2	7E L ND F2	TAGCCTGACACAAGCAAGGG	150	102.9
	7E L ND R2	GCCCGTAATGAAGTCAACCAG		
3	7E ND F1	CTAAAAATGGCCACACAACCA	150	99.3
	7E ND R1	GTCCTTCCAAAAGTGATAGGGATG		
4	7E R ND F1	ACGCGGCCTTTTCATCATTT	300	90.3
	7E R ND R1	GCCCCCTTACTCTGGCATCT		
5	7E R ND F3	GGAAAGCCCAACCAGAATGC	300	102.4
	7E R ND R3	GCATCGAGCGACCCAAGTTA		

Amplicon No.	Primer name	Sequence	Working Conc. (nM)	Efficiency
ChIP at 1.688[7F]				
1	7F L ND F3	CCACTTGGGCTTCAATCGTA	150	101
	7F L ND R3	GCTTGGGGAATACGAGGCA		
2	7F L ND F2	TCTGGTCCTTCGCTGCATTT	150	102.5
	7F L ND R2	GCGACGATATTTGCCTTGGG		
3	7F ND F2	ATCGCCCACCAAGAATCACC	150	101.8
	7F ND R2	TCTTCTTCTCGTGCCTTTGCT		
4	7F R ND F2	ATGCAGGTCGCATTGAGGAA	150	107
	7F R ND R2	CAATGGTCACCCACCCAAGT		
5	7F R ND F4	CGACGTTGGCAGAATAGCAA	300	94.7
	7F R ND R4	CCAAAGGAAAAGCGCACACA		

ChIP at a distance from repeats

Gene name	Primer name	Sequence	Working Conc. (nM)	Efficiency
cin	Cin F2	CTGGTAGCGAAAAGGCCGTA	150	101.8
	Cin R2	CCTCAGCATGTGTTTTCCGC		
opt1 (yin)	Yin F	GGTGATTGCCGAATTCAAGT	300	110.1
	Yin R	ACTAGCATAAGGCTGGCGAA		
CG43689	CG43689 F1	CCCCACAGGTGAGTCATTCC	300	102.1
	CG43689 R1	TGCGGGCTCGTAATAATGCT		
CG1387	CG1387 F1	ACGCTCCATGTCCTTTACGG	300	95.7
	CG1387 R1	CGTTTCGCTTTGCTTTTGCG		
IntS4	IntS4 F1	CACACAGCGGCGTATTTTGT	300	94.9
	IntS4 R1	TGCCCATGAAAGAGTCGGTC		

ChIP at 2L integration site

Amplicon No.	Primer name	Sequence	Working Conc. (nM)	Efficiency
1	SJTG-5kbL flanking F1	AGTTCATCGCCGGATCACTG	150	97.7
	SJTG-5kbL flanking R1	AAAGCGGAGCAGATGGACTT		
2	SJTG-L flanking F1	TCGACTTTGCTCAACACACAA	500	97.3
	SJTG-L flanking R1	ACCCGCATTTCCATTTTGCG		

3	1.688^{3F} insertion	roX1 HincIIR	CTCAATAGCATAAAAAATAGTTGACC	300	91.3
		Sat_LoxP_R	CGACCGTTGCGGCCTTCGTATAGCAT		
3	roX1 insertion	roX1R_706	TCATATCCACTAGCATAAGGCTGGCG	300	103.4
		roX1_LoxP_F	ATCGGTCGGCGGCTTCGTATAATGTA		
4	SJTG-R flanking F1		TACAGTAAGCCACCACCGAT	300	97.6
	SJTG-Rflanking R1		GCGGCAGTCTGTTATCTCTGT	500	
5	SJTG-5kbR flanking F2		GTCTGTAGCAGCAAGCGGTA	150	94.2
	SJTG-5kbR flanking R2		ATCCTGTTGACTGACTGCCG		

B. Primers used for Quantitative RT-PCR primers

Genes on 2L and normalizer

Gene name	Primer name	Sequence	Working Conc. (nM)	Efficiency
CG33128	J1_F2	AACCGACCAGAACCTCATCG	150	98.4
	J1_R2	TCACGGTTCCATTCCAGGTG		
haf	Hap F2	AGCTGAACCTGCTGGATTT	300	95.6
	Hap R2	AGGGTGGACAGCTTTGTTAG		
RFeSP	J2_F3	AAATGATGAACGCCGTGTCG	300	89.2
	J2_R3	GCAGAGCCTTACCCATCGAG		
Eno	J3_F1	ATGTCTTGGACCGCTTCAGT	150	102.1
	J3_R1	GCCCTTTCGATTGGGGTGA		
CG31778	J10_F2	AACCATTCACTGCAGAGGCG	150	95.7
	J10_R2	CCGAAGTCATTGCCCTCAGAA		
Rpl37A	J8_F1	GAGCGCCAAATGGTGACAAA	500	108.6
	J8_R1	CAGGACCAGATGCCCACAAC		
dmn	Dmn_F	GACAAGTTGAGCCGCCTTAC	300	98.5
	Dmn_R	CTTGGTGCTTAGATGACGCA		

REFERENCES

Adachi, Y., Kas, E., and Laemmli, U.K. (1989). Preferential, cooperative binding of DNA topoisomerase II to scaffold-associated regions. EMBO J *8*, 3997-4006.

Akhtar, A., and Becker, P.B. (2000). Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in Drosophila. Mol Cell *5*, 367-375.

Alekseyenko, A.A., Ellison, C.E., Gorchakov, A.A., Zhou, Q., Kaiser, V.B., Toda, N., Walton, Z., Peng, S., Park, P.J., Bachtrog, D.*, et al.* (2013). Conservation and de novo acquisition of dosage compensation on newly evolved sex chromosomes in Drosophila. Genes Dev *27*, 853-858.

Alekseyenko, A.A., Larschan, E., Lai, W.R., Park, P.J., and Kuroda, M.I. (2006). High-resolution ChIP–chip analysis reveals that the Drosophila MSL complex selectively identifies active genes on the male X chromosome. Genes Dev *20*, 848-857.

Alekseyenko, A.A., Peng, S., Larschan, E., Gorchakov, A.A., Lee, O.K., Kharchenko, P., McGrath, S.D., Wang, C.I., Mardis, E.R., Park, P.J.*, et al.* (2008). A sequence motif within chromatin entry sites directs MSL establishment on the Drosophila X chromosome. Cell *134*, 599-609.

Aulner, N., Monod, C., Mandicourt, G., Jullien, D., Cuvier, O., Sall, A., Janssen, S., Laemmli, U.K., and Käs, E. (2002). The AT-Hook Protein D1 Is Essential for Drosophila melanogaster Development and Is Implicated in Position-Effect Variegation. Molecular and Cellular Biology *22*, 1218-1232.

Bachtrog, D., and Charlesworth, B. (2002). Reduced adaptation of a non-recombining neo-Y chromosome. Nature *416*, 323-326.

Bailey, J.A., Carrel, L., Chakravarti, A., and Eichler, E.E. (2000). Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: The Lyon repeat hypothesis. PNAS *97*, 6634-6639.

Bao, S., and Cagan, R. (2006). Fast cloning inverted repeats for RNA interference. RNA *12*, 2020-2024.

Barbash, D.A. (2010). Ninety years of Drosophila melanogaster hybrids. Genetics *186*, 1-8.

Bashaw, G.J., and Baker, B.S. (1997). The regulation of the Drosophila msl-2 gene reveals a function for Sex-lethal in translational control. Cell *89*, 789-798.

Bean, D.M., Heimbach, J., Ficorella, L., Micklem, G., Oliver, S.G., and Favrin, G. (2014). esyN: network building, sharing and publishing. PLoS One *9*, e106035.

Becker, J.S., Nicetto, D., and Zaret, K.S. (2016). H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes. Trends in genetics : TIG *32*, 29-41.

Beckmann, K., Grskovic, M., Gebauer, F., and Hentze, M.W. (2005). A Dual Inhibitory Mechanism Restricts msl-2 mRNA Translation for Dosage Compensation in Drosophila. Cell *122*, 529-540.

Berrios, M., Osheroff, N., and Fisher, P.A. (1985). In situ localization of DNA topoisomerase II, a major polypeptide component of the Drosophila nuclear matrix fraction. PNAS *82*, 4142-4146.

Blattes, R., Monod, C., Susbielle, G., Cuvier, O., Wu, J.H., Hsieh, T.S., Laemmli, U.K., and Kas, E. (2006). Displacement of D1, HP1 and topoisomerase II from satellite heterochromatin by a specific polyamide. EMBO J *25*, 2397-2408.

Boeke, J., Bag, I., Ramaiah, M.J., Vetter, I., Kremmer, E., Pal-Bhadra, M., Bhadra, U., and Imhof, A. (2011). The RNA helicase Rm62 cooperates with SU(VAR)3-9 to re-silence active transcription in Drosophila melanogaster. PLoS One *6*, e20761.

Bone, J.R., and Kuroda, M.I. (1996). Dosage compensation regulatory proteins and the evolution of sex chromosomes in Drosophila. Genetics *144*, 705-713.

Boyd, J.B., Berendes, H.D., and Boyd, H. (1968). Mass Preparation of Nuclei from the Larval Salivary Glands of Drosophila hydei. The Journal of Cell Biology *38*, 369-376.

Cernilogar, F.M., Onorati, M.C., Kothe, G.O., Burroughs, A.M., Parsi, K.M., Breiling, A., Lo Sardo, F., Saxena, A., Miyoshi, K., Siomi, H.*, et al.* (2011). Chromatin-associated RNA interference components contribute to transcriptional regulation in Drosophila. Nature *480*, 391-395.

Charlesworth, B. (1996). The evolution of chromosomal sex determination and dosage compensation. Curr Biol *6*, 149-162.

Chow, J.C., Ciaudo, C., Fazzari, M.J., Mise, N., Servant, N., Glass, J.L., Attreed, M., Avner, P., Wutz, A., Barillot, E.*, et al.* (2010). LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation. Cell *141*, 956-969.

Corona, D.F.V., Siriaco, G., Armstrong, J.A., Snarskaya, N., McClymont, S.A., Scott, M.P., and Tamkun, J.W. (2007). ISWI Regulates Higher-Order Chromatin Structure and Histone H1 Assembly In Vivo. PLOS Biol *5*, e232.

Culbertson, M.R. (1999). RNA surveillance: unforeseen consequences for gene expression, inherited genetic disorders and cancer. Trends in Genetics *15*, 74-80.

de Wit, E., Greil, F., and van Steensel, B. (2005). Genome-wide HP1 binding in Drosophila: developmental plasticity and genomic targeting signals. Genome Res *15*, 1265-1273.

Deal, R.B., and Henikoff, S. (2010). A Simple Method for Gene Expression and Chromatin Profiling of Individual Cell Types within a Tissue. Developmental Cell *18*, 1030-1040.

Deng, H., Zhang, W., Bao, X., Martin, J.N., Girton, J., Johansen, J., and Johansen, K.M. (2005a). The JIL-1 kinase regulates the structure of Drosophila polytene chromosomes. Chromosoma *114*, 173-182.

Deng, X., and Meller, V.H. (2006). roX RNAs are required for increased expression of X-linked genes in Drosophila melanogaster males. Genetics *174*, 1859-1866.

Deng, X., Rattner, B.P., Souter, S., and Meller, V.H. (2005b). The severity of roX1 mutations is predicted by MSL localization on the X chromosome. Mech Dev *122*, 1094-1105.

Deshpande, N., and Meller, V.H. (2014). Sex chromosome evolution: life, death and repetitive DNA. Fly (Austin) *8*, 197-199.

Deuring, R., Fanti, L., Armstrong, J.A., Sarte, M., Papoulas, O., Prestel, M., Daubresse, G., Verardo, M., Moseley, S.L., Berloco, M.*, et al.* (2000). The ISWI chromatin-remodeling protein is required for gene expression and the maintenance of higher order chromatin structure in vivo. Mol Cell *5*, 355-365.

DiBartolomeis, S.M., Tartof, K.D., and Jackson, F.R. (1992). A superfamily of Drosophila satellite related (SR) DNA repeats restricted to the X chromosome euchromatin. Nucleic acids research *20*, 1113-1116.

Disteche, C.M. (2012). Dosage compensation of the sex chromosomes. Annu Rev Genet *46*, 537-560.

Ebert, A., Schotta, G., Lein, S., Kubicek, S., Krauss, V., Jenuwein, T., and Reuter, G. (2004). Su(var) genes regulate the balance between euchromatin and heterochromatin in Drosophila. Genes Dev *18*, 2973-2983.

Ellison, C.E., and Bachtrog, D. (2013). Dosage compensation via transposable element mediated rewiring of a regulatory network. Science *342*, 846-850.

Emilsson, V., Thorleifsson, G., Zhang, B., Leonardson, A.S., Zink, F., Zhu, J., Carlson, S., Helgason, A., Walters, G.B., Gunnarsdottir, S., *et al.* (2008). Genetics of gene expression and its effect on disease. Nature *452*, 423.

Fagegaltier, D., and Baker, B.S. (2004). X chromosome sites autonomously recruit the dosage compensation complex in Drosophila males. PLOS Biol *2*, e341.

Ferree, P.M., and Barbash, D.A. (2009). Species-specific heterochromatin prevents mitotic chromosome segregation to cause hybrid lethality in Drosophila. PLOS Biol *7*, e1000234.

Figueiredo, M.L., Kim, M., Philip, P., Allgardsson, A., Stenberg, P., and Larsson, J. (2014). Non-coding roX RNAs prevent the binding of the MSL-complex to heterochromatic regions. PLoS genetics *10*, e1004865.

Gallach, M. (2014). Recurrent turnover of chromosome-specific satellites in Drosophila. Genome biology and evolution *6*, 1279-1286.

Gebauer, F., Merendino, L., Hentze, M.W., and Valcarcel, J. (1998). The Drosophila splicing regulator sex-lethal directly inhibits translation of male-specific-lethal 2 mRNA. RNA *4*, 142-150.

Gelbart, M.E., and Kuroda, M.I. (2009). Drosophila dosage compensation: a complex voyage to the X chromosome. Development *136*, 1399-1410.

Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. Nature reviews Genetics *10*, 94-108.

Grimaud, C., and Becker, P.B. (2009). The dosage compensation complex shapes the conformation of the X chromosome in Drosophila. Genes Dev *23*, 2490-2495.

Heng, H.H.Q., Goetze, S., Ye, C.J., Liu, G., Stevens, J.B., Bremer, S.W., Wykes, S.M., Bode, J., and Krawetz, S.A. (2004). Chromatin loops are selectively anchored using scaffold/matrix-attachment regions. Journal of Cell Science *117*, 999-1008.

Hilfiker, A., Hilfiker-Kleiner, D., Pannuti, A., and Lucchesi, J.C. (1997). mof, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in Drosophila. EMBO J *16*, 2054-2060.

Hsieh, T., and Brutlag, D. (1979). Sequence and sequence variation within the 1.688 g/cm3 satellite DNA of Drosophila melanogaster. Journal of molecular biology *135*, 465-481.

Jin, Y., Wang, Y., Johansen, J., and Johansen, K.M. (2000). JIL-1, a chromosomal kinase implicated in regulation of chromatin structure, associates with the male specific lethal (MSL) dosage compensation complex. J Cell Biol *149*, 1005-1010.

Joshi, S.S., and Meller, V.H. (2017). Satellite Repeats Identify X Chromatin for Dosage Compensation in Drosophila melanogaster Males. Curr Biol *27*, 1393-1402 e1392.

Kageyama, Y., Mengus, G., Gilfillan, G., Kennedy, H.G., Stuckenholz, C., Kelley, R.L., Becker, P.B., and Kuroda, M.I. (2001). Association and spreading of

the Drosophila dosage compensation complex from a discrete roX1 chromatin entry site. EMBO J *20*, 2236-2245.

Kelley, R.L., and Kuroda, M.I. (2003). The Drosophila roX1 RNA gene can overcome silent chromatin by recruiting the male-specific lethal dosage compensation complex. Genetics *164*, 565-574.

Kelley, R.L., Meller, V.H., Gordadze, P.R., Roman, G., Davis, R.L., and Kuroda, M.I. (1999). Epigenetic spreading of the Drosophila dosage compensation complex from roX RNA genes into flanking chromatin. Cell *98*, 513-522.

Kelley, R.L., Solovyeva, I., Lyman, L.M., Richman, R., Solovyev, V., and Kuroda, M.I. (1995). Expression of msl-2 causes assembly of dosage compensation regulators on the X chromosomes and female lethality in Drosophila. Cell *81*, 867-877.

Kind, J., and Akhtar, A. (2007). Cotranscriptional recruitment of the dosage compensation complex to X-linked target genes. Genes Dev *21*, 2030-2040.

Kind, J., Vaquerizas, J.M., Gebhardt, P., Gentzel, M., Luscombe, N.M., Bertone, P., and Akhtar, A. (2008). Genome-wide analysis reveals MOF as a key regulator of dosage compensation and gene expression in Drosophila. Cell *133*, 813-828.

Koya, S.K., and Meller, V.H. (2011). roX RNAs and Genome Regulation in Drosophila Melanogaster. Prog Mol Subcell Biol *51*, 147-160.

Koya, S.K., and Meller, V.H. (2015). Modulation of Heterochromatin by Male Specific Lethal Proteins and roX RNA in Drosophila melanogaster Males. PLoS One *10*, e0140259.

Kuhn, G.C., Kuttler, H., Moreira-Filho, O., and Heslop-Harrison, J.S. (2012). The 1.688 repetitive DNA of Drosophila: concerted evolution at different genomic scales and association with genes. Molecular biology and evolution *29*, 7-11.

Kuzu, G., Kaye, E.G., Chery, J., Siggers, T., Yang, L., Dobson, J.R., Boor, S., Bliss, J., Liu, W., Jogl, G.*, et al.* (2016). Expansion of GA Dinucleotide Repeats Increases the Density of CLAMP Binding Sites on the X-Chromosome to Promote Drosophila Dosage Compensation. PLoS genetics *12*, e1006120.

Lamb, J., Crawford, E.D., Peck, D., Modell, J.W., Blat, I.C., Wrobel, M.J., Lerner, J., Brunet, J.-P., Subramanian, A., Ross, K.N.*, et al.* (2006). The Connectivity Map: Using Gene-Expression Signatures to Connect Small Molecules, Genes, and Disease. Science *313*, 1929-1935.

Larschan, E., Alekseyenko, A.A., Gortchakov, A.A., Peng, S., Li, B., Yang, P., Workman, J.L., Park, P.J., and Kuroda, M.I. (2007). MSL Complex Is Attracted to Genes Marked by H3K36 Trimethylation Using a Sequence-Independent Mechanism. Molecular Cell *28*, 121-133.

Larschan, E., Bishop, E.P., Kharchenko, P.V., Core, L.J., Lis, J.T., Park, P.J., and Kuroda, M.I. (2011). X chromosome dosage compensation via enhanced transcriptional elongation in Drosophila. Nature *471*, 115-118.

Lee, J.T. (2009). Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. Genes Dev *23*, 1831-1842.

Lim, C.K., and Kelley, R.L. (2013). The Drosophila over compensating males gene genetically inhibits dosage compensation in males. PLoS One *8*, e60450.

Lipardi, C., and Paterson, B.M. (2009). Identification of an RNA-dependent RNA polymerase in Drosophila involved in RNAi and transposon suppression. PNAS *106*, 15645-15650.

Liu, L.P., Ni, J.Q., Shi, Y.D., Oakeley, E.J., and Sun, F.L. (2005). Sex-specific role of Drosophila melanogaster HP1 in regulating chromatin structure and gene transcription. Nat Genet *37*, 1361-1366.

Lohe, A.R., Hilliker, A.J., and Roberts, P.A. (1993). Mapping simple repeated DNA sequences in heterochromatin of Drosophila melanogaster. Genetics *134*, 1149-1174.

Lucchesi, J.C., and Kuroda, M.I. (2015). Dosage compensation in Drosophila. Cold Spring Harb Perspect Biol *7*.

Lyon, M.F. (1998). X-chromosome inactivation: a repeat hypothesis. Cytogenet Cell Genet *80*, 133-137.

Lyon, M.F. (2006). Do LINEs have a role in X-chromosome inactivation? Journal of biomedicine & biotechnology *2006*, 59746.

McElroy, K.A., Jung, Y.L., Zee, B.M., Wang, C.I., Park, P.J., and Kuroda, M.I. (2017). upSET, the Drosophila homologue of SET3, Is Required for

Viability and the Proper Balance of Active and Repressive Chromatin Marks. G3 (Bethesda) *7*, 625-635.

Meller, V.H. (2003). Initiation of dosage compensation in Drosophila embryos depends on expression of the roX RNAs. Mech Dev *120*, 759-767.

Meller, V.H., Fisher, P.A., and Berrios, M. (1995). Intranuclear distribution of DNA topoisomerase II and chromatin. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *3*, 255-260.

Meller, V.H., Joshi, S.S., and Deshpande, N. (2015). Modulation of Chromatin by Noncoding RNA. Annu Rev Genet *49*, 673-695.

Meller, V.H., and Rattner, B.P. (2002). The roX genes encode redundant male-specific lethal transcripts required for targeting of the MSL complex. EMBO J *21*, 1084-1091.

Mendjan, S., Taipale, M., Kind, J., Holz, H., Gebhardt, P., Schelder, M., Vermeulen, M., Buscaino, A., Duncan, K., Mueller, J., *et al.* (2006). Nuclear pore components are involved in the transcriptional regulation of dosage compensation in Drosophila. Mol Cell *21*, 811-823.

Menon, D.U., Coarfa, C., Xiao, W., Gunaratne, P.H., and Meller, V.H. (2014). siRNAs from an X-linked satellite repeat promote X-chromosome recognition in Drosophila melanogaster. PNAS *111*, 16460-16465.

Menon, D.U., and Meller, V.H. (2009). Imprinting of the Y Chromosome Influences Dosage Compensation in roX1 roX2 Drosophila melanogaster</em&gt. Genetics *183*, 811.

Menon, D.U., and Meller, V.H. (2012). A role for siRNA in X-chromosome dosage compensation in Drosophila melanogaster. Genetics *191*, 1023-1028.

Meyer, B.J. (2000). Sex in the wormcounting and compensating X-chromosome dose. Trends Genet *16*, 247-253.

Moazed, D. (2009). Small RNAs in transcriptional gene silencing and genome defence. Nature *457*, 413-420.

Monfort, M., and Furlong, E.E. (2015). Furlong Lab EMBL TFBS contribution.

Morales, V., Straub, T., Neumann, M.F., Mengus, G., Akhtar, A., and Becker, P.B. (2004). Functional integration of the histone acetyltransferase MOF into the dosage compensation complex. EMBO J *23*, 2258-2268.

Park, Y., Kelley, R.L., Oh, H., Kuroda, M.I., and Meller, V.H. (2002). Extent of Chromatin Spreading Determined by roX RNA Recruitment of MSL Proteins. Science *298*, 1620.

Partridge, J.F., Borgstrom, B., and Allshire, R.C. (2000). Distinct protein interaction domains and protein spreading in a complex centromere. Genes Dev *14*, 783-791.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic acids research *29*, e45.

Quinn, J.J., Ilik, I.A., Qu, K., Georgiev, P., Chu, C., Akhtar, A., and Chang, H.Y. (2014). Revealing long noncoding RNA architecture and functions using domain-specific chromatin isolation by RNA purification. Nat Biotechnol *32*, 933-940.

Ramirez, F., Lingg, T., Toscano, S., Lam, K.C., Georgiev, P., Chung, H.-R., Lajoie, B., de Wit, E., Zhan, Y., de Laat, W., *et al.* (2015). High-affinity sites form an interaction network to facilitate spreading of the MSL complex across the X chromosome in Drosophila. Molecular cell *60*, 146-162.

Rastelli, L., Richman, R., and Kuroda, M.I. (1995). The dosage compensation regulators MLE, MSL-1 and MSL-2 are interdependent since early embryogenesis in Drosophila. Mech Dev *53*, 223-233.

Reik, W., and Walter, J. (2001). Genomic imprinting: parental influence on the genome. Nature Reviews Genetics *2*, 21.

Rice, W.R. (1996). Evolution of the Y sex chromosome in animals. Bioscience *Vol. 46*, 331-343.

Riddle, N.C., Minoda, A., Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Tolstorukov, M.Y., Gorchakov, A.A., Jaffe, J.D., Kennedy, C., Linder-Basso, D., *et al.* (2011). Plasticity in patterns of histone modifications and chromosomal proteins in Drosophila heterochromatin. Genome Res *21*, 147-163.

Samata, M., and Akhtar, A. (2018). Dosage Compensation of the X Chromosome: A Complex Epigenetic Assignment Involving Chromatin Regulators and Long Noncoding RNAs. Annual Review of Biochemistry *87*, null.

Scott, M.J., Pan, L.L., Cleland, S.B., Knox, A.L., and Heinrich, J. (2000). MSL1 plays a central role in assembly of the MSL complex, essential for dosage compensation in Drosophila. EMBO J *19*, 144-155.

Seum, C., Pauli, D., Delattre, M., Jaquet, Y., Spierer, A., and Spierer, P. (2002). Isolation of Su(var)3-7 mutations by homologous recombination in Drosophila melanogaster. Genetics *161*, 1125-1136.

Shogren-Knaak, M., Ishii, H., Sun, J.M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science *311*, 844-847.

Smith, E.R., Allis, C.D., and Lucchesi, J.C. (2001). Linking global histone acetylation to the transcription enhancement of X-chromosomal genes in Drosophila males. J Biol Chem *276*, 31483-31486.

Smith, E.R., Pannuti, A., Gu, W., Steurnagel, A., Cook, R.G., Allis, C.D., and Lucchesi, J.C. (2000). The drosophila MSL complex acetylates histone H4 at lysine 16, a chromatin modification linked to dosage compensation. Mol Cell Biol *20*, 312-318.

Soruco, M.M., Chery, J., Bishop, E.P., Siggers, T., Tolstorukov, M.Y., Leydon, A.R., Sugden, A.U., Goebel, K., Feng, J., Xia, P.*, et al.* (2013). The CLAMP protein links the MSL complex to the X chromosome during Drosophila dosage compensation. Genes Dev *27*, 1551-1556.

Spierer, A., Begeot, F., Spierer, P., and Delattre, M. (2008). SU(VAR)3-7 links heterochromatin and dosage compensation in Drosophila. PLoS genetics *4*, e1000066.

Spierer, A., Seum, C., Delattre, M., and Spierer, P. (2005). Loss of the modifiers of variegation Su(var)3-7 or HP1 impacts male X polytene chromosome morphology and dosage compensation. Journal of Cell Science *118*, 5047.

Stark, C., Breitkreutz, B.J., Reguly, T., Boucher, L., Breitkreutz, A., and Tyers, M. (2006). BioGRID: a general repository for interaction datasets. Nucleic acids research *34*, D535-539.

Steinemann, M., Steinemann, S., and Turner, B.M. (1996). Evolution of dosage compensation. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *4*, 185-190.

Straub, T., Grimaud, C., Gilfillan, G.D., Mitterweger, A., and Becker, P.B. (2008). The chromosomal high-affinity binding sites for the Drosophila dosage compensation complex. PLoS genetics *4*, e1000302.

Straub, T., Zabel, A., Gilfillan, G.D., Feller, C., and Becker, P.B. (2013). Different chromatin interfaces of the Drosophila dosage compensation complex revealed by high-shear ChIP-seq. Genome Res *23*, 473-485.

Sugiyama, T., Cam, H., Verdel, A., Moazed, D., and Grewal, S.I. (2005). RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. PNAS *102*, 152-157.

Sural, T.H., Peng, S., Li, B., Workman, J.L., Park, P.J., and Kuroda, M.I. (2008). The MSL3 chromodomain directs a key targeting step for dosage

compensation of the Drosophila melanogaster X chromosome. Nat Struct Mol Biol *15*, 1318-1325.

Swaminathan, A., Gajan, A., and Pile, L.A. (2012). Epigenetic regulation of transcription in Drosophila. Front Biosci (Landmark Ed) *17*, 909-937.

Tamura, K., Subramanian, S., and Kumar, S. (2004). Temporal patterns of fruit fly (Drosophila) evolution revealed by mutation clocks. Molecular biology and evolution *21*, 36-44.

Urban, J., Kuzu, G., Bowman, S., Scruggs, B., Henriques, T., Kingston, R., Adelman, K., Tolstorukov, M., and Larschan, E. (2017). Enhanced chromatin accessibility of the dosage compensated Drosophila male X-chromosome requires the CLAMP zinc finger protein. PLoS One *12*, e0186855.

Usakin, L., Abad, J., Vagin, V.V., de Pablos, B., Villasante, A., and Gvozdev, V.A. (2007). Transcription of the 1.688 satellite DNA family is under the control of RNA interference machinery in Drosophila melanogaster ovaries. Genetics *176*, 1343-1349.

Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I., and Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. Science *303*, 672-676.

Villa, R., Schauer, T., Smialowski, P., Straub, T., and Becker, P.B. (2016). PionX sites mark the X chromosome for dosage compensation. Nature *537*, 244.

Wang, C.I., Alekseyenko, A.A., LeRoy, G., Elia, A.E., Gorchakov, A.A., Britton, L.M., Elledge, S.J., Kharchenko, P.V., Garcia, B.A., and Kuroda, M.I.

(2013). Chromatin proteins captured by ChIP-mass spectrometry are linked to dosage compensation in Drosophila. Nat Struct Mol Biol *20*, 202-209.

Wang, Y., Zhang, W., Jin, Y., Johansen, J., and Johansen, K.M. (2001). The JIL-1 tandem kinase mediates histone H3 phosphorylation and is required for maintenance of chromatin structure in Drosophila. Cell *105*, 433-443.

Waring, G.L., and Pollack, J.C. (1987). Cloning and characterization of a dispersed, multicopy, X chromosome sequence in Drosophila melanogaster. PNAS *84*, 2843-2847.

Weinmann, A.S., Plevy, S.E., and Smale, S.T. (1999). Rapid and Selective Remodeling of a Positioned Nucleosome during the Induction of IL-12 p40 Transcription. Immunity *11*, 665-675.

Wu, L., Zee, Barry M., Wang, Y., Garcia, Benjamin A., and Dou, Y. (2011). The RING Finger Protein MSL2 in the MOF Complex Is an E3 Ubiquitin Ligase for H2B K34 and Is Involved in Crosstalk with H3 K4 and K79 Methylation. Molecular Cell *43*, 132-144.

Zhang, K., Mosch, K., Fischle, W., and Grewal, S.I. (2008). Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. Nat Struct Mol Biol *15*, 381-388.

Zhang, W., Deng, H., Bao, X., Lerach, S., Girton, J., Johansen, J., and Johansen, K.M. (2006). The JIL-1 histone H3S10 kinase regulates dimethyl H3K9 modifications and heterochromatic spreading in Drosophila. Development *133*, 229-235.

Zhou, R., Czech, B., Brennecke, J., Sachidanandam, R., Wohlschlegel, J.A., Perrimon, N., and Hannon, G.J. (2009). Processing of Drosophila endo-siRNAs depends on a specific Loquacious isoform. RNA *15*, 1886-1895.

ABSTRACT

ROLE OF SIRNA PATHWAY IN EPIGENETIC MODIFICATIONS OF THE DROSOPHILA MELANOGASTER X CHROMOSOME

by

NIKITA DESHPANDE

August 2018

Advisor: Dr. Victoria H. Meller

Major: Biological Sciences

Degree: Doctor of Philosophy

Eukaryotic genomes are organized into large domains of coordinated regulation. The role of small RNAs in formation of these domains is largely unexplored. An extraordinary example of domain-wide regulation is X chromosome compensation in *Drosophila melanogaster* males. This process occurs by hypertranscription of genes on the single male X chromosome. Extensive research in this field has shown that the Male Specific Lethal (MSL) complex binds X-linked genes and modifies chromatin to increase expression. The components of this complex, and their actions on chromatin, are well studied. In contrast, the mechanism that results in exclusive recruitment to the X chromosome is not understood. Our research focuses on the process by which male flies selectively modulate expression from their single X chromosome. Prior studies in the laboratory have found that the siRNAs produced from repetitive sequences on the X chromosome and the repeat DNA itself, participates in dosage compensation in flies. Interestingly, the siRNA pathway contributes to X-

localization of the MSL complex. The basis of enhanced localization is unknown, and no RNAi components have been found to interact directly with the MSL complex. This suggests that siRNA influences X-recognition by an indirect and novel mechanism. I found evidence that chromatin around these repeats is modulated by the siRNA pathway. I demonstrated that FLAG-tagged Argonaute2 protein localizes at these repeats. We show that numerous Agonaute2-interacting proteins show evidence of participation in compensation. One of these, Su(var)3-9, deposits H3K9me2 in and near the repeats. When a repeat-containing transgene is inserted on an autosome, H3K9me2 is enriched in surrounding chromatin, an effect that is enhanced by ectopic production of cognate siRNA. In accord with the idea that these repeats contribute to recruitment of dosage compensation, genes as much as 100 kb from the autosomal insertion increase in expression upon expression of ectopic siRNA. My studies demonstrate that chromatin around a group of X-enriched sequences is modulated by siRNA, and supports the idea that siRNA contributes to the elevated expression that characterizes the compensated male X chromosome. This study advances our understanding of the mechanism of X recognition by showing a direct relationship between siRNA-directed chromatin modification and a class of repetitive elements that helps mark the X chromosome.

AUTOBIOGRAPHICAL STATEMENT

NIKITA DESHPANDE

Educational Qualification:

2011-2018: Doctor of Philosophy (Biological Sciences), Wayne State University, Detroit, MI, USA

2009-2011: Master of Science (Medical Biotechnology), The Maharaja Sayajirao University, Vadodara, India

2006-2009: Bachelor of Science (Microbiology), University of Mumbai, Mumbai, India

Honors and Awards:

2016 WSU Graduate Student Professional Travel Award

2016 Summer Dissertation Fellowship, WSU

2015 Poster Presentation Award at the 6th Annual Graduate Research Day, WSU

2014 Excellence in Teaching Award, WSU

2013 Graduate Enhancement Award, WSU

Publications:

Deshpande N, Meller VH (*submitted*, Genetics). Chromatin at X-linked repeats that guide dosage compensation in *Drosophila melanogaster* is modulated by the siRNA pathway

Meller VH, Joshi SS, **Deshpande N** (2015). Modulation of Chromatin by Noncoding RNA. Annual review of genetics 2015; 49:673-95.

Deshpande N, Meller VH (2015). Sex chromosome evolution: life, death and repetitive DNA. Fly; 8:197-9.